THE LANCET
Microbe

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Rosado J, Pelleau S, Cockram C, et al. Multiplex assays for the identification of serological signatures of SARS-CoV-2 infection: an antibody-based diagnostic and machine learning study. *Lancet Microbe* 2020; published online Dec 19. https://doi.org/10.1016/S2666-5247(20)30197-X.

Multiplex assays for the identification of serological signatures of SARS-CoV-2 infection: an antibody-based diagnostic and machine learning study

Appendix and Supplementary Methods

Jason ROSADO, Stéphane PELLEAU, Charlotte COCKRAM, Sarah Hélène MERKLING, Narimane NEKKAB, Caroline DEMERET, Annalisa MEOLA, Solen KERNEIS, Benjamin TERRIER, Samira FAFI-KREMER, Jerome de SEZE, Timothée BRUEL, François DEJARDIN, Stéphane PETRES, Rhea LONGLEY, Arnaud FONTANET, Marija BACKOVIC, Ivo MUELLER, Michael WHITE*

***Correspondence to Dr Michael White: michael.white@pasteur.fr**

Table of Contents

1. Multiplex serological assay and classification algorithms

1.1. Serological assays

In a first step, four proteins derived from SARS-CoV-2 Spike were included in the assay. This includes SARS-CoV-2 trimeric Spike ectodomain (S^{tri}) and its receptor-binding domain (RBD) produced as recombinant proteins in mammalian cells in the Structural Virology Unit at Institut Pasteur, while S1 (cat# REC31806) and S2 (cat# REC31807) subunits were purchased from Native Antigen, Oxford, UK. S^{tri} and RBD were designed based on the viral genome sequence of the SARS-CoV-2 strain France/IDF0372/2020, obtained from the GISAID database (accession number EPI_ISL_406596). The synthetic genes, codon-optimized for protein expression in mammalian cells, were ordered from GenScript (Leiden, Netherlands) and cloned in pcDNA3.1(+) vector as follows: the RBD, residues 331-519, and the entire S ectodomain (residues 1-1208). The RBD construct included an exogenous signal peptide of a human kappa light chain (METDTLLLWVLLLWVPGSTG) to ensure efficient protein secretion into the media. The S ectodomain construct was engineered, as reported before to have the stabilizing double proline mutation (KV986-987 to PP986-987) and the foldon domain at the C-terminus that allows the S to trimerize (YIPEAPRDGQAYVRKDGEWVLLSTFL) resembling the native S state on the virion [1]. Both constructs contained a Strep (WSHPQFEK), an octa-histidine, and an Avi tag (GLNDIFEAQKIEWHE) at the C-terminus for affinity purification. Protein expression was done by transient transfection of mammalian HEK293 free style cells. Proteins were then purified from supernatants on a Streptactin column (IBA Biosciences, IBA GmbH, Göttingen, Germany) followed by size exclusion purification on Superdex 200 column using standard chromatography protocols.

In a second step, eight proteins were added to the assay. Recombinant SARS-CoV-2 nucleoprotein (NP) was expressed in *E. coli* in the Production and Purification of Recombinant Proteins Technological Platform at Institut Pasteur. Two SARS-CoV-2 antigens were purchased from Native Antigen, Oxford, UK: RBD (cat# REC31831-20) and NP (cat# REC31812-100). The His-tagged SARS-CoV-2 N protein was bacterially expressed in E. coli BL21 (DE3) at Institut Pasteur and purified as a soluble dimeric protein by affinity purification using a Ni-NTA Protino column (Macherey Nagel) and gel filtration using a HiLoad 16/60 Superdex 200 pg column (GE Healthcare). Additional antigens for seasonal coronaviruses 229E NP (cat# REC31758-100) and NL63 NP (cat# REC31759-100), influenza A (cat# FLU-H1N1-HA-100), adenovirus type 40 (cat# NAT41552-100) and rubella (cat# REC31651-100) were purchased from Native Antigen. All proteins were coupled to magnetic beads as described elsewhere [2]. The mass of proteins coupled on beads were optimized to generate a log-linear standard curve with a pool of positive serum prepared from RT-qPCR-confirmed SARS-CoV-2 patients.

In total, we optimized a 12-plex assay able to detect antibody responses against seven SARS-CoV-2 antigens (two nucleoproteins constructs, five spike), one nucleoprotein for each seasonal coronavirus NL63 and 229E, and three antigens from other viruses (Influenza A H1N1, adenovirus type 40, rubella) for which a large part of the population is expected to be seropositive due to vaccination or natural infection and hence serve as internal controls (Appendix Table 1).

The assay was performed in black, 96 well, non-binding microtiter plate (cat#655090; Greiner Bio-One, Germany). Briefly 50 µL of protein-conjugated magnetic beads (500/region/µL) and 50 µL of diluted serum were mixed and incubated for 30 min at room temperature on a plate shaker. All dilutions were made in phosphate buffered saline containing 1% bovine serum albumin and 0·05% (v/v) Tween-20 (denoted as PBT), and all samples were run in singlicate. Following incubation, the magnetic beads were separated using magnetic plate separator (Luminex®, Austin, Texas, USA) for 60 seconds and washed three times with 100 µL of PBT. The washed magnetic beads were incubated for 15 minutes with detector secondary antibody at room temperature on a plate shaker. The magnetic beads were separated and washed three times with 100 µL of PBT and finally resuspended in 100 µL of PBT. Two separate assays were used for measuring IgG and IgM antibodies. For IgM measurements, serum samples were diluted 1/200, and R-Phycoerythrin (R-PE) -conjugated Donkey Anti-Human IgM (cat#709-116-073; JacksonImmunoResearch, UK) antibody was used as secondary antibody at 1/400 dilution. For IgG, serum samples were diluted 1/100, and R-Phycoerythrin (R-PE) -conjugated Donkey Anti-Human IgG (cat#709-116-098; JacksonImmunoResearch, UK) antibody was used as secondary antibody at 1/120 dilution.

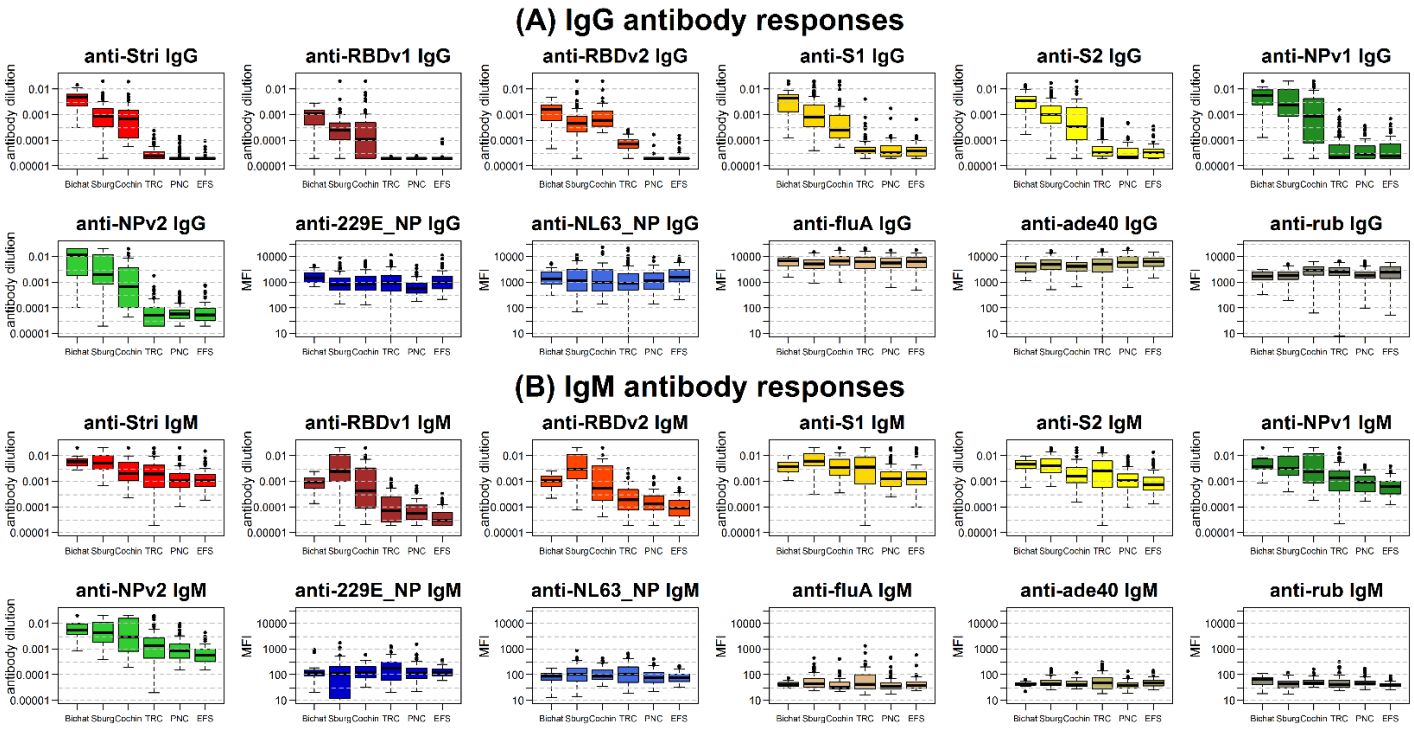
On each plate, two blanks (only beads, no serum) were included as well as a standard curve prepared from two-fold serial dilutions (1:50 to 1:25600) of a pool of positive controls. Plates were read using a Luminex® MAGPIX® system and the median fluorescence intensity (MFI) was used for analysis. A 5-parameter logistic curve was used to convert MFI to antibody dilution, relative to the standard curve performed on the same plate to account for inter-assay variations. The multiplex immunoassay was validated by checking that the MFI obtained were well correlated with those obtained in monoplex (only one conjugated bead type per well). For non-SARS-CoV-2 antigens, MFI data was used for the analysis.

The antigens included in the multiplex assay were selected with the aim of providing multiple measurements of the antibody response directed towards the Spike protein and the Nucleoprotein. Factors affecting antigen selection included availability and timeliness. Thus for example, at the time when this assay was being developed our team had access to the nucleoprotein of two seasonal alpha-coronaviruses (NL63, 229E), but we did not yet have access to antigens from the two seasonal beta-coronaviruses (OC43, HKU1).

Appendix Table 1: List of antigens included in the multiplex serological assay.

category	short name	recombinant antigen	expression system	supplier, catalog number
SARS-CoV-2	S^{tri}	SARS-CoV-2 Trimeric Spike protein	HEK293	Institut Pasteur, Paris
SARS-CoV-2	RBD_{v1}	SARS-CoV-2 Spike Glycoprotein (S1) RBD	CHO	Native Antigen, REC31831-100
SARS-CoV-2	RBD_{v2}	SARS-CoV-2 Spike Glycoprotein (S1) RBD	HEK293	Institut Pasteur, Paris
SARS-CoV-2	S1	SARS-CoV-2 Spike Glycoprotein (S1)	HEK293	Native Antigen, REC31806-100
SARS-CoV-2	S2	SARS-CoV-2 Spike Glycoprotein (S2)	HEK293	Native Antigen, REC31807-100
SARS-CoV-2	NP_{v1}	SARS-CoV-2 Nucleoprotein	*E. coli*	Institut Pasteur, Paris
SARS-CoV-2	NP_{v2}	SARS-CoV-2 Nucleoprotein	*E. coli*	Native Antigen, REC31812-100
seasonal coronavirus	229E-NP	Human Coronavirus 229E Nucleoprotein	*E. coli*	Native Antigen, REC31758-100
seasonal coronavirus	NL63-NP	Human Coronavirus NL63 Nucleoprotein	*E. coli*	Native Antigen, REC31759-100
internal controls	FluA	Influenza virus H1N1 haemagglutinin recombinant antigen	HEK293	Native Antigen, FLU-H1N1-HA-100
internal controls	Ade40	Adenovirus type 40 Hexon (capside)	HEK293	Native Antigen, NAT41552-100
internal controls	Rub	Rubella virus-like particles (spike glycoprotein E1, spike glycoprotein E2 and Capsid protein)	HEK293	Native Antigen, REC31651-100

An overview of the output of the 12 plex assay for IgG and IgM antibody responses is provided in Appendix Figure 1.



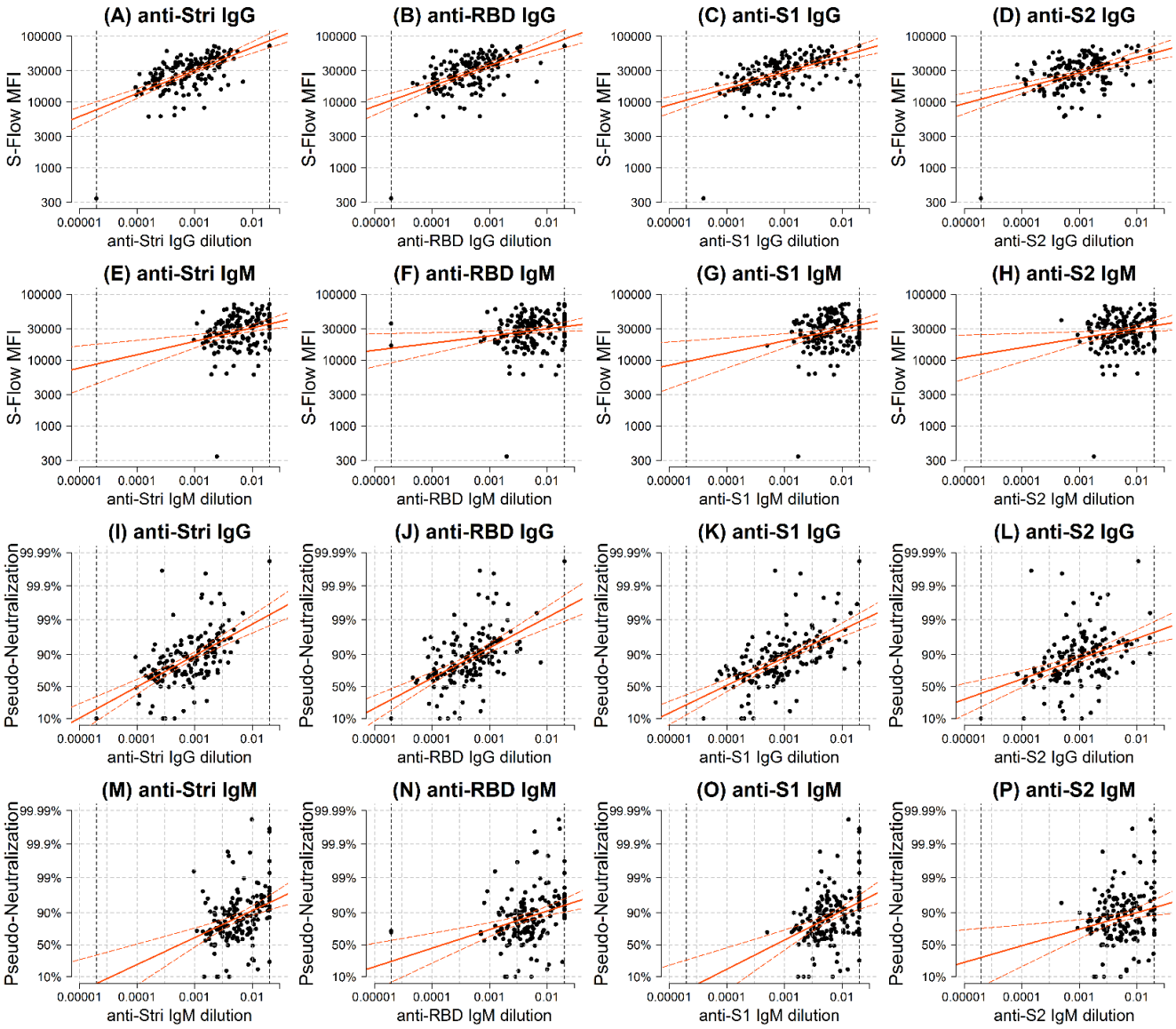
Appendix Figure 1: Anti-SARS-CoV-2 antibody responses. (A) Measured IgG antibody dilutions or medium fluorescence intensity (MFI) in serum samples with previously confirmed RT-qPCR infection from patients in Hôpital Bichat (*n* = 34), health care workers from Strasbourg (*n* = 162), and Hôpital Cochin (*n* = 63). Negative control samples from Thailand (*n* = 68), Peru (*n* = 90), and French donors (*n* = 177) were also tested. **(B)** Measured IgM antibody dilutions or MFI in serum or plasma samples. **(C)** Receiver Operating Characteristic (ROC) curve for IgG antibodies obtained by varying the cutoff for seropositivity. Colours correspond to those shown in part A. **(D)** ROC curve for IgM antibodies obtained by varying the cutoff for seropositivity. **(E)** Area under the ROC curve for individual biomarkers. **(F)** Spearman correlation between measured antibody responses.

1.2. Comparison with other immunoassays

The multiplex Luminex assay was validated by comparison with data from other serological assays on matched samples. The samples used for this validation were the 162 samples from healthcare workers in Strasbourg hospitals with RT-qPCR confirmed SARS-CoV-2 infection. Two other serological assays were used: (i) S-Flow; and (ii) pseudo-neutralization [3,4]. In the S-Flow assay, a flow-cytometry based assay that measures antibodies binding to the spike protein (S) (GenBank: QHD43416.1) expressed at the surface of 293T cells (ATCC® CRL-3216™). Two parameters can be calculated with the S-Flow assay: the first is the percentage of cells having captured antibodies, defining the seropositivity. The second is the mean fluorescence intensity (MFI) of this binding, which provides a quantitative measurement of the amount of antibodies and their efficacy. Samples were also tested for neutralization activity at a single dilution of 1:100 using a viral pseudotype-based assay. Briefly, single cycle lentiviral pseudotypes coated with the S protein and encoding for a luciferase reporter gene were preincubated with the serum to be tested at a dilution of 1:100, and added to 293T-ACE2 target cells (Addgene

Plasmid #1786). The luciferase signal was measured after 48 h. The percentage of neutralization was calculated by comparing the signal obtained with each serum to the signal generated by control negative sera.

Appendix Figure 2 shows the association between IgG and IgM antibody levels measured on the Luminex assay (S^{tri}, RBD, S1, S2) and antibody responses measured using the S-Flow and pseudo-neutralization assays. There were statistically significant associations between all 16 pairwise comparisons of the biomarkers studied. This indicates that our multiplex Luminex assay provides qualitatively and quantitatively similar output to other immunoassays.



Appendix Figure 2; Association between immunoassays in samples from Strasbourg healthcare workers. (A-H) IgG and IgM antibody levels to trimeric spike, RBD, S1 and S2 measured using a multiplex Luminex assay and S-Flow MFI. **(I-P)** IgG and IgM antibody levels to trimeric spike, RBD, S1 and S2 measured using a multiplex Luminex assay and pseudo-neutralization activity. Each point represents a sample from one of 162 healthcare workers from hospitals in Strasbourg. The solid orange line represents the fit of a linear regression model, with the 95% confidence intervals represented as orange dashed lines.

1.3. Pairwise comparison of classification performance of biomarkers

The classification performance of a biomarker can be assessed on its potential to correctly classify positive and negative samples, where classification will depend on the choice of cutoff. The classification performance of two biomarkers can be statistically compared using McNemar's test. For cutoffs corresponding to a high specificity target, Appendix Table 2 provides a pairwise comparison of the classification performance of 14 SARS-CoV-2 biomarkers (7 IgG and 7 IgM). The best individual biomarker is anti-S^{tri} IgG antibody levels, which provide significantly better classification than all other biomarkers.

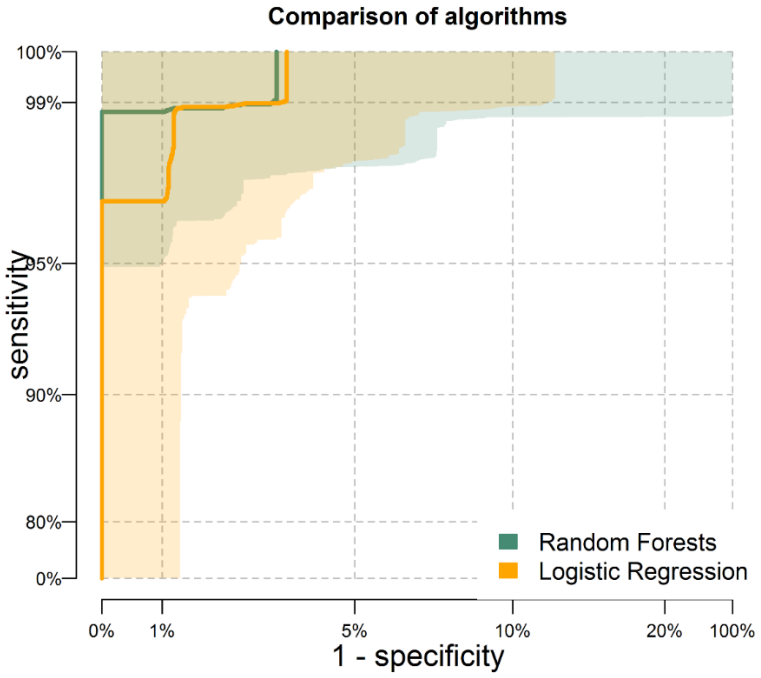
Appendix Table 2: Comparison of classification performance between biomarkers for a high specificity target (>99%). Pairwise comparisons are made using McNemar's test. The above diagonal element shows the odds ratio with 95% confidence intervals. Odds ratio > 1 indicates that biomarker indicated by the row has better classification than the biomarker indicated by the column. The corresponding element below the diagonal presents the P value.

		S^{tri} IgG	RBD_{v1} IgG	RBD_{v2} IgG	S1 IgG	S2 IgG	NP_{v1} IgG	NP_{v2} IgG	S^{tri} IgM	RBD_{v1} IgM	RBD_{v2} IgM	S1 IgM	S2 IgM	NP_{v1} IgM	NP_{v2} IgM
S^{tri}	**IgG**		13·75 (5·1, 52·3)	5·3 (2·7, 11·7)	Inf (30·1, Inf)	87 (15·2, 3475)	6·8 (3·4, 15·5)	9 (4·1, 23)	220 (39·1, 8729)	20·3 (7·6, 76·1)	25·5 (9·7, 95·4)	107 (29·3, 889)	Inf (64·0, Inf)	100·5 (27·5, 835·6)	65 (21·9, 317·9)
RBD_{v1}	**IgG**	$< 10^{-10}$		0·76 (0·4, 1·3)	5·77 (3·2, 11·3)	2·35 (1·5, 3·9)	1·04 (0·6, 1·8)	1·2 (0·7, 2·1)	43 (16·5, 159·6)	2·18 (1·3, 3·8)	3·35 (2·0, 5·8)	41·25 (15·8, 153·2)	Inf (50·2, Inf)	25·67 (11·5, 71·0)	21·1 (10·0, 53·5)
RBD_{v2}	**IgG**	$3{\cdot}4 \times 10^{-8}$	0·36		15 (6·1, 47·5)	2·87 (1·76,4·83)	1·23 (0·8, 1·9)	1·33 (0·9, 2·0)	45 (17·3, 166·9)	2·06 (1·3, 3·3)	3·0 (1·9, 4·9)	29·17 (13·1, 80·6)	66 (22·3, 322·8)	32·2 (13·5, 100·5)	25·83 (11·6, 71·5)
S1	**IgG**	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$		0·4 (0·22,0·7)	0·25 (0·14, 0·41)	0·28 (0·17, 0·45)	14·25 (6·99, 33·8)	0·48 (0·31, 0·73)	0·75 (0·5,1·12)	8·62 (4·8, 16·7)	32·25 (12·3, 120·2)	9·6 (5, 20·7)	7·08 (3·9, 13·8)
S2	**IgG**	$< 10^{-10}$	$2{\cdot}2 \times 10^{-4}$	$5{\cdot}7 \times 10^{-6}$	$9{\cdot}0 \times 10^{-4}$		0·47 (0·29, 0·73)	0·48 (0·3, 0·77)	20 (9·5, 50·7)	0·83 (0·55, 1·26)	1·27 (0·84, 1·92)	11·5 (6·4, 22·8)	26·33 (11·8, 72·8)	11·27 (6·1, 23·2)	8·57 (4·9, 16·2)
NP_{v1}	**IgG**	$< 10^{-10}$	1·0	0·39	$7{\cdot}0 \times 10^{-10}$	$5{\cdot}9 \times 10^{-4}$		1·57 (0·56, 4·78)	56·67 (19·1, 277·5)	1·89 (1·18, 3·11)	2·92 (1·81, 4·85)	54·33 (18·27, 266·2)	187 (33·2, 7425·1)	30·4 (12·8, 94·9)	29 (12·2, 90·7)
NP_{v2}	**IgG**	$< 10^{-10}$	0·59	0·21	$1{\cdot}0 \times 10^{-8}$	$1{\cdot}6 \times 10^{-3}$	0·48		28·17 (12·7, 77·8)	1·7 (1·06, 2·76)	2·62 (1·6, 4·3)	32·2 (13·5, 100·5)	183 (32·4, 7267)	18·88 (9·3, 44·5)	16·11 (8·3, 35·9)
S^{tri}	**IgM**	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$		0·01 (0, 0·05)	0·02 (0, 0·06)	0·53 (0·2, 1·34)	3·11 (1·43, 7·49)	0·44 (0·23, 0·82)	0·36 (0·18, 0·66)
RBD_{v1}	**IgM**	$< 10^{-10}$	$2{\cdot}6 \times 10^{-3}$	$7{\cdot}7 \times 10^{-4}$	$4{\cdot}7 \times 10^{-4}$	0·42	$7{\cdot}3 \times 10^{-3}$	0·026	$< 10^{-10}$		Inf (5·21, Inf)	68·5 (18·6, 571·3)	162 (28·7, 6437)	41·67 (13·9, 204·8)	24 (10,75·26)
RBD_{v2}	**IgM**	$< 10^{-10}$	$4{\cdot}3 \times 10^{-7}$	$1{\cdot}3 \times 10^{-7}$	0·17	0·28	$2{\cdot}2 \times 10^{-6}$	$1{\cdot}7 \times 10^{-5}$	$< 10^{-10}$	$9{\cdot}5 \times 10^{-7}$		39 (13·0, 191·8)	47·67 (16·0, 233·9)	26·25 (9·95, 98·13)	14·43 (6·75, 36·8)
S1	**IgM**	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	0·21	$< 10^{-10}$	$< 10^{-10}$		7·5 (2·6, 29·3)	0·59 (0·.32, 1·08)	0·47 (0·25, 0·85)
S2	**IgM**	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$2{\cdot}6 \times 10^{-3}$	$< 10^{-10}$	$< 10^{-10}$	$6{\cdot}2 \times 10^{-6}$		0·13 (0·05, 0·31)	0·12 (0·04, 0·27)
NP_{v1}	**IgM**	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$7{\cdot}8 \times 10^{-3}$	$< 10^{-10}$	$< 10^{-10}$	0.092	$1{\cdot}8 \times 10^{-8}$		0·12 (0, 0·93)
NP_{v2}	**IgM**	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$< 10^{-10}$	$4{\cdot}6 \times 10^{-4}$	$< 10^{-10}$	$< 10^{-10}$	0.01	$3{\cdot}0 \times 10^{-10}$	0·039	

1.4. Selection of classification algorithms

In previous work on the analysis of data from a multiplex assay of *Plasmodium vivax* malaria antibody responses, we tested the performance of a number of different classification algorithms including logistic regression, linear discriminant analyses, quadratic discriminant analyses, decision trees, and random forests [5]. This analysis found random forests algorithms to have the highest classification performance of the tested algorithms. On this basis, we selected random forests algorithms to classify previous SARS-CoV-2 infection using multiplex serological data.

To further test this, we compared the classification performance of a random forests algorithm with a logistic regression classifier. The data from positive and negative samples as detailed in Table 1 for assessing classification performance. 1000-fold repeat cross-validation was implemented with splitting of the data into disjoint training sets (2/3 of the data) and testing sets (1/3 of the data). Appendix Figure 3 shows cross-validated Receiver Operating Characteristic (ROC) curves for Random Forests and Logistic Regression classifiers using six measurements of antibody responses (S^{tri} IgG; RBD_{v2} IgG; NP_{v1} IgG; S2 IgG; RBD_{v1} IgM; NP_{v1} IgM). For all values of specificity, the random forests classifier provides equal or better sensitivity than the logistic regression classifier. At a fixed value of 99% specificity, the logistic regression classifier had 96·6% (95% CI: 88·2%, 100·0%) sensitivity, whereas the random forests classifier had 98·9% (95% cI: 97·1%, 99·6%) sensitivity.



Appendix Figure 3: Comparison of random forests and logistic regression classifiers. ROC curves were estimated using 1000-fold cross-validation with training (2/3 of samples) and testing (1/3 of samples) data sets. Shaded regions depict regions of 95% uncertainty.

1.5. Results of multiplex classification with random forests algorithm

The classification performance of a diagnostic test can be assessed against a number of sensitivity and specificity targets. Here we consider three targets in detail:

- **high sensitivity target:** sensitivity > 99% is enforced, and subject to these constraints specificity is maximised.
- **balanced target:** sensitivity and specificity are equally weighted, so that a combination that minimises |sensitivity – specificity| is selected.
- **high specificity target:** specificity > 99% is enforced, and subject to these constraints sensitivity is maximised.

For these three targets, Appendix Table 3 shows the sensitivity and specificity for the 14 individual biomarkers, plus multiplex combinations up to size six. Multiplex combinations were selected on their potential to optimize the high specificity target. For example, anti-S^{tri} IgG antibody levels provided the greatest sensitivity for combinations of size 1. Therefore, for combinations of size 2, anti-S^{tri} IgG antibody levels were automatically included, and all other possible biomarkers were included in combinations. The pairwise combination that maximized sensitivity was anti-S^{tri} IgG antibody levels and anti-RBD^{v2} IgG antibody levels. In general, the optimal multiplex combination of size *n* was selected by starting with optimal multiplex combination of size *n* – 1 and adding possible remaining biomarkers.

Note that the routine described above does not necessarily guarantee that the optimal combination of size *n* will be selected. To account for this possibility, we tested all possible combinations of IgG antibody levels to the seven SARS-CoV-2 antigens. This resulted in a total of $2^7 - 1 = 127$ combinations. The output of all of these combinations against a range of targets is summarized in detail in a spreadsheet provided in the GitHub repository associated with the publication (https://github.com/MWhite-InstitutPasteur/SARS_CoV_2_SeroDX_phase2/blob/master/SARSCoV2_IgG_antigen_combination.xlsx).

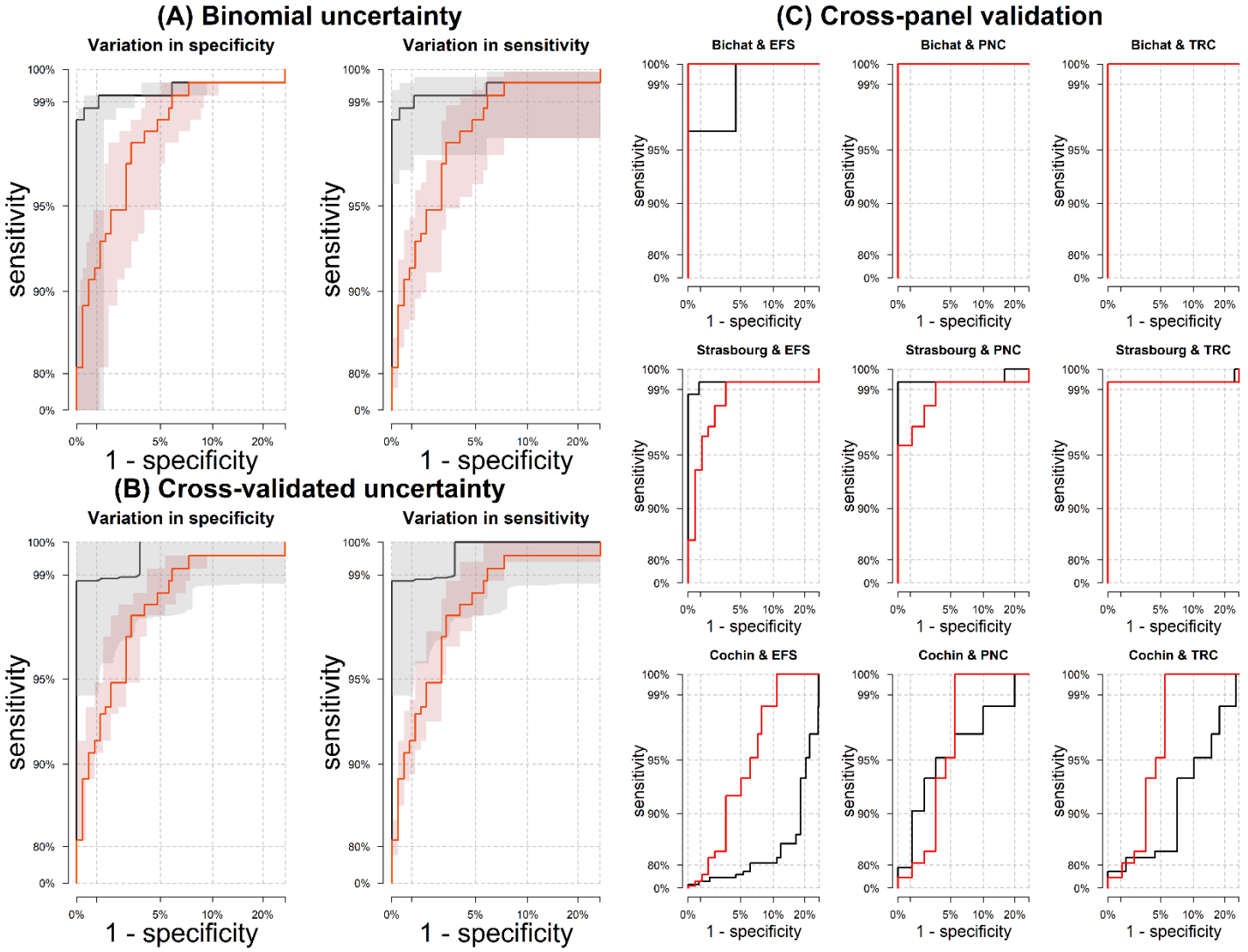
Appendix Table 3: Sensitivity and specificity targets for single biomarkers and multiplex combinations. 95% binomial confidence intervals were calculated using Wilson's method. Antigen combinations were selected to optimize sensitivity for the high specificity target, i.e. the highest sensitivity while enforcing specificity > 99%.

biomarker	high sensitivity target (sensitivity > 99%)		balanced target (sensitivity ~ specificity)		high specificity target (specificity > 99%)	
	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity
IgG antibody dilution						
anti-S^{tri}	99·2% (97·1%, 99·8%)	94·0% (91·0%, 96·1%)	97·1% (94·4%, 98·6%)	97·0% (94·6%, 98·4%)	91·6% (87·5%, 94·5%)	99·1% (97·4%, 99·7%)
anti-RBD_{v1}	100·0% (98·5%, 100·0%)	0·0% (0·0%, 1·1%)	89·2% (84·8%, 92·5%)	97·5% (95·3%, 98·7%)	88·0% (83·5%, 91·5%)	99·1% (97·4%, 99·7%)
anti-RBD_{v2}	99·2% (97·1%, 99·8%)	0% (0%, 1·1%)	91·2% (87·1%, 94·1%)	91·4% (87·9%, 93·9%)	78·1% (72·6%, 82·8%)	99·1% (97·4%, 99·7%)
anti-S1	99·2% (97·1%, 99·8%)	75·9% (71·0%, 80·2%)	91·6% (87·5%, 94·5%)	91·7% (88·2%, 94·2%)	50·6% (44·5%, 56·7%)	99·1% (97·4%, 99·7%)
anti-S2	99·2% (97·1%, 99·8%)	22·6% (18·5%, 27·4%)	90·7% (86·6%, 93·7%)	91·4% (87·9%, 93·9%)	66·9% (60·9%, 72·5%)	99·1% (97·4%, 99·7%)
anti-NP_{v1}	100·0% (98·5%, 100·0%)	0% (0%, 1·1%)	85·7% (80·8%, 89·5%)	85·8% (81·7%, 89·1%)	73·7% (67·9%, 78·8%)	99·1% (97·4%, 99·7%)
anti-NP_{v2}	100·0% (98·5%, 100·0%)	0% (0%, 1·1%)	86·9% (82·1%, 90·5%)	86·7% (82·7%, 89·9%)	72·9% (67·1%, 78·0%)	99·1% (97·4%, 99·7%)
IgM antibody dilution						
anti-S^{tri}	99·6% (97·8%, 99·9%)	9·2% (6·6%, 12·8%)	74·9% (69·2%, 79·9%)	74·9% (70·0%, 79·2%)	0% (0%, 1·5%)	100% (98·9%, 100%)
anti-RBD_{v1}	99·2% (97·1%, 99·8%)	24·0% (19·7%, 28·8%)	87·3% (82·6%, 90·8%)	87·1% (83·1%, 90·3%)	64·5% (58·4%, 70·2%)	99·3% (97·8%, 99·8%)
anti-RBD_{v2}	99·2% (97·1%, 99·8%)	28·4% (23·9%, 33·5%)	84·9% (79·9%, 88·8%)	84·5% (80·2%, 88·0%)	57·0% (50·8%, 62·9%)	99·3% (97·8%, 99·8%)
anti-S1	99·6% (97·8%, 99·9%)	7·4% (5·0%, 10·7%)	69·3% (63·4%, 74·7%)	69·4% (64·2%, 74·1%)	0% (0%, 1·5%)	100% (98·9%, 100%)
anti-S2	99·2% (97·8%, 99·8%)	8·5% (6·0%, 12·0%)	65·7% (59·7%, 71·3%)	65·7% (60·5%, 70·6%)	0% (0%, 1·5%)	100% (98·9%, 100%)
anti-NP_{v1}	99·2% (97·1%, 99·8%)	17·3% (13·7%, 21·8%)	73·3% (67·5%, 78·4%)	73·4% (68·5%, 77·9%)	0% (0%, 1·5%)	100% (98·9%, 100%)
anti-NP_{v1}	99·2% (97·1%, 99·8%)	5·5% (3·6%, 8·5%)	73·7% (67·9%, 78·8%)	73·8% (68·8%, 78·2%)	0% (0%, 1·5%)	100% (98·9%, 100%)
Multiplex combinations						
S^{tri} IgG + RBD_{v2} IgG	99·2% (97·2%, 99·8%)	95·8% (93·1%, 97·5%)	97·6% (94·9%, 98·9%)	97·6% (95·4%, 98·8%)	95·6% (92·3%, 97·5%)	99·1% (97·4%, 99·7%)
S^{tri} IgG + RBD_{v2} IgG + NP_{v1} IgG	99·2% (97·2%, 99·8%)	96·6% (94·1%, 98·1%)	98·0% (95·4%, 99·1%)	98·1% (96·1%, 99·1%)	98·0% (95·4%, 99·1%)	99·1% (97·4%, 99·7%)
S^{tri} IgG + RBD_{v2} IgG + NP_{v1} IgG + S2 IgG	99·2% (97·2%, 99·8%)	96·0% (93·3%, 97·6%)	98·8% (96·5%, 99·6%)	98·8% (96·9%, 99·5%)	98·4% (96·0%, 99·4%)	99·1% (97·4%, 99·7%)
S^{tri} IgG + RBD_{v2} IgG + NP_{v1} IgG + S2 IgG + RBD_{v1} IgM	99·2% (97·2%, 99·8%)	98·5% (96·6%, 99·4%)	98·8% (96·5%, 99·6%)	98·9% (97·1%, 99·6%)	98·8% (96·5%, 99·6%)	99·3% (97·6%, 99·8%)
S^{tri} IgG + RBD_{v2} IgG + NP_{v1} IgG + S2 IgG + RBD_{v1} IgM + NP_{v1} IgM	99·2% (97·2%, 99·8%)	98·9% (97·1%, 99·6%)	98·8% (96·5%, 99·6%)	98·9% (97·1%, 99·6%)	98·8% (96·5%, 99·6%)	99·3% (97·6%, 99·8%)

1.6. Uncertainty in classification performance

Uncertainty in sensitivity and specificity can be quantified in three ways: (i) binomial confidence intervals calculated using Wilson's method; (ii) 1000-fold repeat cross-validation with a training set comprising 2/3 of the data and a disjoint testing set comprising 1/3 of the data; (iii) cross-panel validation with algorithms trained and tested on disjoint panels of data.

For a single antigen (S^{tri}) IgG assay, and a six antigen multiplex assay with random forests classifier (Stri IgG + RBDv2 IgG + NPv1 IgG + S2 IgG + RBDv1 IgM + NPv1 IgM), Appendix Figure 4 provides a detailed overview of the uncertainty in assay sensitivity and specificity.



Appendix Figure 4: Quantification of uncertainty for serological classification. Results are shown for a single antigen (S^{tri}) IgG assay in red, and a six antigen multiplex classifier in black. **(A)** Uncertainty estimated using Wilson's binomial method applied to data from all samples. Uncertainty in specificity at fixed sensitivity, and variation in sensitivity at fixed specificity are shown separately. **(B)** Uncertainty estimated using 1000-fold cross-validation with training (2/3 of samples) and testing (1/3 of samples) data sets. Uncertainty in specificity at fixed sensitivity, and variation in sensitivity at fixed specificity are shown separately. **(C)** Cross-panel validation. The title of each plot denotes the panels that were used for testing, while the other panels were used for training.

2. Antibody kinetics models

2.1. Prior longitudinal data on long-term antibody responses to coronaviruses

There are limited available longitudinal data on SARS-CoV-2 antibody kinetics, and no data from long-term follow-up (as of July 2020). However, there are a number of published studies on the long-term antibody kinetics to other coronaviruses, most notably Severe Acute Respiratory Syndrome coronavirus (SARS-CoV). Here we review some of the available published data, and describe how this can be used to provide prior information for modelling SARS-CoV-2 antibody kinetics.

Appendix Table 4 summarises some of the published data on the long-term antibody kinetics to a number of coronaviruses: SARS-CoV, human seasonal coronavirus 229E, and Middle East Respiratory Syndrome coronavirus (MERS-CoV) [6-12]. From the extracted time series, we estimated two summary statistics characterizing the long-term antibody response: the half-life of the long-lived component of the antibody response, and the percentage reduction in antibody response after one year. The half-life of the long-lived component of the antibody response was estimated by fitting a linear regression model to measurements of (log) antibody response taken greater than six months after symptom onset. The percentage reduction in antibody response after one year was estimated based on the reduction from the peak measured antibody response to the estimated antibody level at one year. Although a wide range of assays from ELISA to micro-neutralisation were used in the reviewed studies, in this simple and approximate analysis we did not attempt to account for assay dependent effects, except to subtract background antibody levels where necessary.

Based on the estimated summary statistics, we assume that the long-term IgG antibody kinetics can be characterized as having a half-life of d_l = 400 days with a 60% reduction after one year. In terms of the parameters of the mathematical model of antibody kinetics, this corresponds to prior estimates of $c_l = \log(2)/d_l = 0.0017$ and $\rho \sim 0.9$. For sensitivity analyses, we also considered scenarios where d_l = 200 days and d_l = 800 days. For IgM antibody kinetics, we assumed d_l = 100 days and $\rho \sim 0{\cdot}9$. For sensitivity analyses, we also considered scenarios where d_l = 50 days and d_l = 200 days.

A prior estimate of the half-life of IgG molecules of 21 days is assumed [13]. A prior estimate of the half-life of IgM molecules of 10 days is assumed [14]. Prior estimates for the short-lived component of the antibody response (half-life = 3·5 days) are consistent with data from several sources [15-19].

Appendix Table 4: Prior data on the duration of antibody responses to coronaviruses. Data from longitudinal studies on measured antibody levels to SARS coronavirus, seasonal coronavirus 229E, and MERS coronavirus. For each study, the time series describing the antibody kinetics was extracted. The half-life of the long-lived component of the antibody response was estimated using measurements of antibody response measured after 6 months from symptom onset – the subset of the data used for this calculation is indicated in bold below. The percentage reduction in antibodies after one year is estimated based on the reduction from the peak measured response to the estimate antibody level at year.

study												**half-life (days)**	**1 year reduction**
SARS-CoV; Wu *et al.* [6]													
time (days)	180	365	730	1095									
IgG	**0·96**	**0·638**	**0·516**	**0·249**								510	
SARS-CoV; Mo *et al*. [7]													
time (days)	7	15	30	60	90	180	270	360	450	540	720		
IgG	0·01	1·86	2·36	2·83	2·81	**2·73**	**2·38**	**1·91**	**1·42**	**1·00**	**0·80**	181	60%
IgM	0·01	1·13	1·80	1·30	0·69	0·06	0·01						100%
Nab	0·01	1·99	2·74	2·51	2·26	**2·06**	**1·83**	**1·56**	**1·24**	**0·96**	**0·78**	277	69%
SARS-CoV; Cao *et al.* [8]													
time (days)	30	120	210	300	480	720	900	1080					
IgG	196	244	**114**	**112**	**64**	**36**	**33**	**28**				394	61%
Nab	1034	1254	**836**	**773**	**960**	**99**	**32**	**32**				154	33%
SARS-CoV; Liu *et al.* [9]													
time (days)	30	120	210	300	480	720							
IgG	185	201	**115**	**125**	**65**	**32**						254	49%
SARS-CoV; Tang *et al*. [10]													
time (days)	24	120	210	300	480	720	900	1080	1600	2160			
IgG	305	252	**128**	**170**	**66**	**31**	**36**	**33**	**6.9**	**6.0**		400	57%
seasonal coronavirus 229E; [11]													
time (days)	0	21	84	364									
IgG	2·45	3·18	**2·62**	**2·51**								191	91%
IgA	2·61	3·04	**2·80**	**2·66**								150	87%
Nab	1·43	9·84	**5·46**	**2·19**								116	91%
MERS CoV; Choe *et al*. [12]													
time (days)	15	90	200	300	400								
IgG (S1)	1·39	2·53	**1·63**	**1·56**	**1·47**							915	50%

2.2. Mathematical model of antibody kinetics

SARS-CoV-2 antibody kinetics are described using a previously published mathematical model of the immunological processes underlying the generation and waning of antibody responses following infection or vaccination [15]. The existing model is adapted to account for the frequent data available in the first weeks of infection.

$$\frac{dB}{dt} = -bB,$$
$$\frac{dP_s}{dt} = \rho B - c_s P_s,$$
$$\frac{dP_l}{dt} = (1-\rho)B - c_l P_l,$$
$$\frac{dA}{dt} = gP_s + gP_l - rA$$

where B denotes B lymphocytes, $b > 0$ is the rate of differentiation of B lymphocytes into antibody secreting plasma cells, P_s denotes short-lived plasma cells, P_l denotes long-lived plasma cells, ρ $(0 \leq \rho \leq 1)$ is the proportion of plasma cells that are short-lived, $g > 0$ is the rate of generation of antibodies (IgG or IgM) from plasma cells, and $r > 0$ is the rate of decay of antibody molecules. Assuming $B(0) = B_0$, $P_s(0) = P_l(0) = 0$ and $A(0) = A_{bg}$, these equations can be solved analytically to give:

$$A(t) = A_{bg} + gB_0\left(\frac{(\rho c_l + (1-\rho)c_s - \delta)e^{-b(t-\delta)}}{(c_s-\delta)(c_l-\delta)(r-b)} + \frac{(\rho c_l + (1-\rho)c_s - r)e^{-r(t-\delta)}}{(c_s-r)(c_l-r)(b-r)} + \frac{\rho e^{-c_s(t-\delta)}}{(c_s-r)(c_s-b)} + \frac{(1-\rho)e^{-c_l(t-\delta)}}{(c_l-r)(c_l-b)}\right)$$

δ is the time after symptom onset when antibody levels start to increase. B_0 is the number of B cells, and g is the rate at which they secrete antibodies. As g and B_0 are not both identifiable without detailed and invasive experiments (e.g. bone marrow aspirates to measure antigen-specific plasma cells), we estimate $\beta = gB_0$. If r is the decay rate of antibody molecules, then we define $d_a = \frac{\log(2)}{r}$ to be the half-life of antibody molecules. Similarly, we define $d_b = \frac{\log(2)}{b}$, $d_s = \frac{\log(2)}{c_s}$ and $d_l = \frac{\log(2)}{c_l}$.

2.3. Methodology for statistical inference

The model was fitted to longitudinal antibody level measurements from all participants. Mixed effects methods were used to capture the natural variation in antibody kinetics between individual participants, whilst estimating the average value and variance of the immune parameters across the entire population of individuals. The models were fitted in a Bayesian framework using Markov Chain Monte Carlo (MCMC) methods. Mixed effects methods allow individual-level parameters to be estimated for each participant separately, with these individual-level (or mixed effects) parameters being drawn from global distributions. For example, for each participant n the half-life of the short-lived ASCs may be estimated as d_s^n (an individual-level parameter). These N estimates of the local parameters d_s^n will be drawn from a probability distribution. A log-Normal distribution is suitable as it has positive support on [0, ∞). Thus we have $log(d_s^n) \sim N(\mu_s, \sigma_s^2)$. The mean d_s and the variance Σ_s^2 of the estimates of d_s^n are given by $d_s = e^{\mu_s + \frac{\sigma_s^2}{2}}$ and $\Sigma_s^2 = \left(e^{\sigma_s^2} - 1\right)e^{2\mu_s + \sigma_s^2}$.

Model likelihood

For individual *n* we have data on observed antibody levels $A^n = \{a_1, \dots, a_J\}$ at times $T^n = \{t_1, \dots, t_J\}$. We denote $D^n = (A^n, T^n)$ to be the vector of data for individual *n*. For individual *n*, the parameters $\theta^n = \left(A_{bg}^n, \delta^n, B_0^n, \delta^n, d_s^n, d_l^n, d_a^n, \rho^n\right)$ are estimated. The model predicted antibody levels will be $\{A(t_1), A(t_2), \dots, A(t_J)\}$. We assume log-Normally distributed measurement error such that the difference between $log(a_j)$ and $log\left(A(t_j)\right)$ is Normally distributed with variance σ_{obs}^2. For model predicted antibody levels $A(t_j)$ the data likelihood for individual *n* is given by

$$L_{mod}^n(\theta^n|D^n) = \prod_{j \in J} \frac{e^{-\frac{\left(log(a_j) - log\left(A(t_j)\right)\right)^2}{2\sigma_{obs}^2}}}{a_j \sigma_{obs}\sqrt{2\pi}}$$

Mixed effects likelihood

As described above, for each individual there are eight parameters to be estimated. The mixed effects likelihood can be written as follows:

$$L_{mix}^n(\theta^n|D^n) = \left(\frac{e^{-\frac{\left(log\left(A_{bg}^n\right) - \mu_A\right)^2}{2\sigma_A^2}}}{\sqrt{2\pi}A_{bg}^n\sigma_A}\right)\left(\frac{e^{-\frac{(log(\delta^n) - \mu_\delta)^2}{2\sigma_\delta^2}}}{\sqrt{2\pi}\delta^n\sigma_\delta}\right)\left(\frac{e^{-\frac{(log(\beta^n) - \mu_\beta)^2}{2\sigma_\beta^2}}}{\sqrt{2\pi}\beta^n\sigma_\beta}\right)\left(\frac{e^{-\frac{(log(d_b^n) - \mu_b)^2}{2\sigma_b^2}}}{\sqrt{2\pi}d_b^n\sigma_b}\right)\left(\frac{e^{-\frac{(log(d_s^n) - \mu_s)^2}{2\sigma_s^2}}}{\sqrt{2\pi}d_s^n\sigma_s}\right)$$

$$\left(\frac{e^{-\frac{(log(d_l^n) - \mu_l)^2}{2\sigma_l^2}}}{\sqrt{2\pi}d_l^n\sigma_l}\right)\left(\frac{e^{-\frac{(log(d_a^n) - \mu_a)^2}{2\sigma_a^2}}}{\sqrt{2\pi}d_a^n\sigma_a}\right)\left(\frac{e^{-\frac{\left(log\left(\frac{\rho^n}{1-\rho^n}\right) - \mu_\rho\right)^2}{2\sigma_\rho^2}}}{\sqrt{2\pi}\rho^n(1-\rho^n)\sigma_\rho}\right) L_{mod}^n(\theta^n|D^n)$$

As the proportion of the ASCs that are long-lived must be bounded by 0 and 1, the individual-level parameters ρ^n are assumed to be drawn from logit-Normal distributions.

Total model likelihood

Denote $D = \{D^1, \dots, D^N\}$ to be the vector of data for all *N* participants. We denote $\theta^n = \left(\mu_A, \sigma_A, \mu_\delta, \sigma_\delta, \mu_\beta, \sigma_\beta, \mu_b, \sigma_b, \mu_s, \sigma_s, \mu_l, \sigma_l, \mu_a, \sigma_a, \mu_\rho, \sigma_\rho, \theta^1, \dots, \theta^N\right)$ to be the combined vector of population-level parameters and individual-level parameters to be estimated. The total likelihood is obtained by multiplying the likelihood for each participant

$$L_{total}(\theta|D) = \prod_{n \in N} L^n_{mix}(\theta^n|D^n)$$

Markov Chain Monte Carlo parameter update

The model was fitted to the data using Markov Chain Monte Carlo (MCMC) methods. A three stage parameter update regimen was utilised with a Metropolis-within-Gibbs sampler with sequential updating of individual-level parameters, population-level parameters, and observational variance parameters. A' indicates an attempted update.

1. **Individual-level parameter Metropolis-Hastings update.** For each participant *n*:
   - Update local parameters: $\theta^{n\prime} = \left(A^{n}_{bg}{}', \delta^{n\prime}, \beta^{n\prime}, d^{n\prime}_{b}, d^{n\prime}_{s}, d^{n\prime}_{l}, d^{n\prime}_{a}, \rho^{n\prime}\right)$
   - Calculate updated mixed effects likelihood $L^n_{mix}(\theta^{n\prime}|D^n)$
   - Accept the parameter update with probability $min\left(1, \frac{L^n_{mix}\left(\theta^{n\prime}|D^n\right)}{L^n_{mix}(\theta^n|D^n)}\right)$

2. **Population-level parameter Gibbs update.**
   - For each of the $i \in \{A, \delta, \beta, b, s, l, a\}$ we obtain new estimates of the population level parameters $\mu_i{}'$ and $\tau_i' = \frac{1}{\sigma_i{}'}$ as follows:

$$\mu_i' \sim N\left(\frac{\tau_{i,0}\mu_{i,0} + \tau_i \sum_N log(y_i^n)}{\tau_{i,0} + \tau_i N}, \frac{1}{\sqrt{\tau_{i,0} + \tau_i N}}\right)$$

$$\tau_i' \sim \Gamma\left(\frac{N}{2} + k, \left(\frac{1}{\theta} + \frac{1}{2}\sum_N (log(y_i^n) - \mu_i{}')^2\right)^{-1}\right)$$

   where $\mu_{i,0}$ and $\tau_{i,0}$ parameterise the Normal prior distribution on the mean, and *k* and θ parameterise the Gamma prior distribution on the precision (inverse standard deviation). y_i^n is the individual-level parameter *i* in individual *n*.

   - For ρ we obtain new estimates of the population level parameters $\mu_\rho{}'$ and $\tau_\rho' = \frac{1}{\sigma_\rho{}'}$ as follows:

$$\mu_\rho' \sim N\left(\frac{\tau_{\rho,0}\mu_{\rho,0} + \tau_\rho \sum_N logit(\rho^n)}{\tau_{\rho,0} + \tau_\rho N}, \frac{1}{\sqrt{\tau_{\rho,0} + \tau_\rho N}}\right)$$

$$\tau_\rho' \sim \Gamma\left(\frac{N}{2} + k, \left(\frac{1}{\theta} + \frac{1}{2}\sum_N \left(logit(\rho^n) - \mu_\rho{}'\right)^2\right)^{-1}\right)$$

   where $\mu_{\rho,0}$ and $\tau_{\rho,0}$ parameterise the Normal prior distribution on the mean, and *k* and θ parameterise the Gamma prior distribution on the precision (inverse standard deviation).

3. **Observational variance parameter Metropolis-Hastings Update.**
   - Update the observational variance parameter $\sigma_{obs}{}'$

- Calculate updated total likelihood $L_{total}(\theta'|D)$and the updated prior probability density $P(\theta')$
- Accept the parameter update with probability $min\left(1, \frac{L_{total}(\theta\prime|D)P(\theta\prime)}{L_{total}(\theta|D)P(\theta)}\right)$

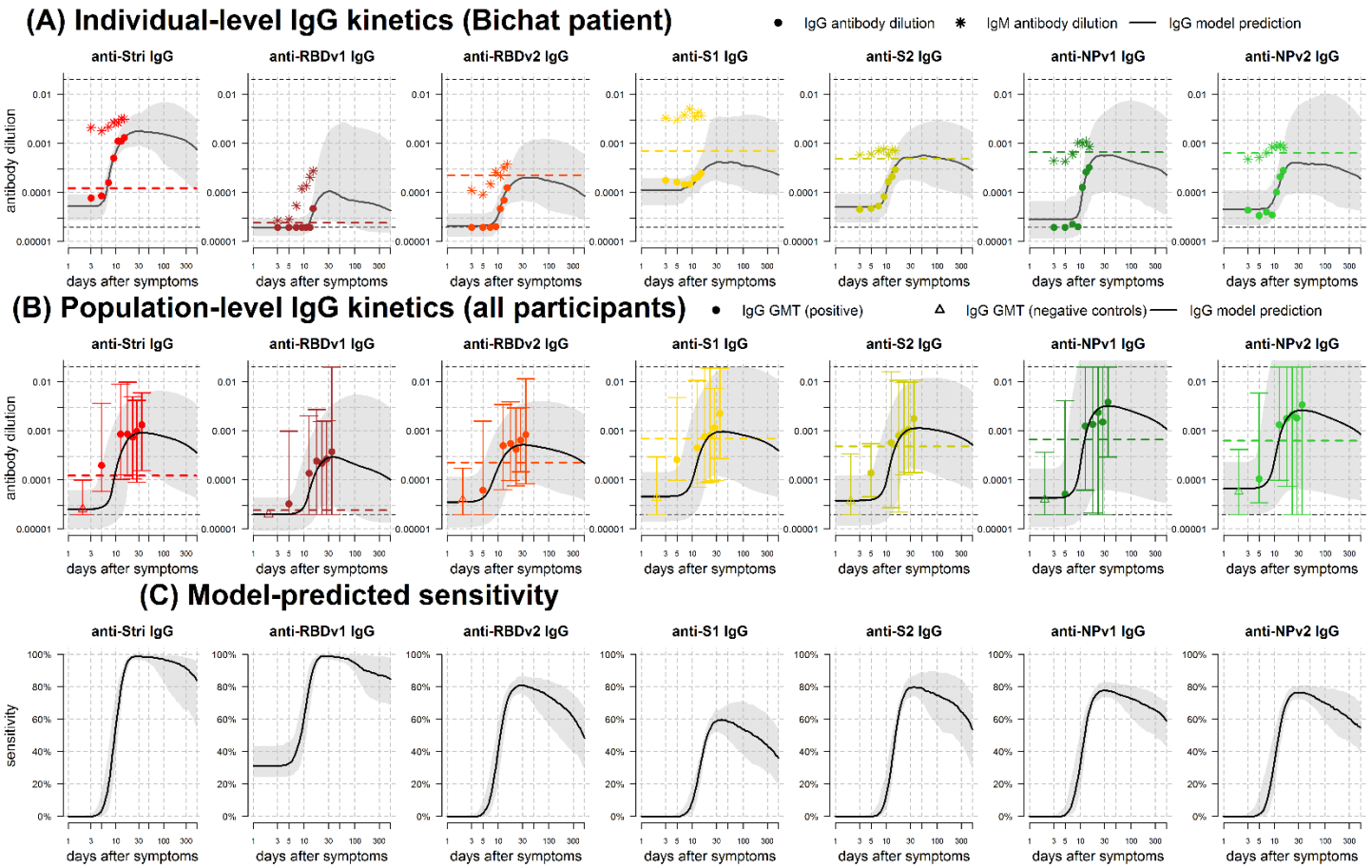
The MCMC algorithm was implemented in C++ complied in Microsoft Visual Studio. The covariance of the multivariate-Normal proposal distributions for Metropolis-Hastings updates were adaptively tuned using the estimated posterior distributions during a burn-in phase of 1 million MCMC iterations. The magnitude of the proposed step size was calibrated using a Robbins-Munro algorithm to ensure an acceptance rate of approximately 23%. The total number of MCMC iterations was 10,000,000. The effective number of iterations was calculated using the effectiveSize routine in the R library coda and the effective size was checked to be > 1,000 for all parameters.

2.4. Fit of antibody kinetic model to data.

For all seven SARS-CoV-2 antigens, we fit the mathematical models of antibody kinetics to data on IgG antibody levels. We did not attempt to fit the model to data on IgM antibody levels because of higher levels of noise in the IgM assay, and a lack of prior data on the long-term kinetics of IgM antibodies. Prior and posterior parameter estimates with 95% credible intervals are presented in Appendix Table 5.

Appendix Figure 5 shows data from a patient from Hôpital Bichat with frequent longitudinal sampling. The data and model indicate that the antibody response is in a rising phase between 5 and 30 days after symptom onset. The seroconversion time depends on the seropositivity cutoff. For the cutoffs shown, seroconversion occurs for anti-S^{tri} IgG, anti-RBD_{v1} IgG and anti-S2 IgG, but not for anti-RBD_{v2} IgG, anti-S1 IgG, anti-NP_{v1} IgG and anti-NP_{v2} IgG.

For all 215 individuals with RT-qPCR SARS-CoV-2 infection, Appendix Figure 5B shows the model predicted IgG antibody response to SARS-CoV-2. For all antigens, we predict a bi-phasic pattern of waning with a first rapid phase between one and three months after symptom onset, followed by a slower rate of waning. The percentage reduction in antibody level after one year was mostly determined by prior information and estimated to be 47% (95% CrI: 18%, 90%) for anti-S^{tri} IgG antibodies, with comparable estimates for other antigens (Appendix Figure 5B & Appendix Table 6). Sensitivity was assessed using the seropositivity cutoff based on a high specificity target (>99%). For all antigens considered, we predict that there will be a reduction in sensitivity over time, although there is a large degree of uncertainty (Appendix Figure 5C).



Appendix Figure 5: IgG antibody kinetics. (A) Measured IgG antibody dilutions, shown as points, from a patient in Hôpital Bichat followed longitudinally. Posterior median model predictions of IgG antibody dilution are shown as black lines, with 95% credible intervals in grey. The coloured dashed line represents the cutoff for IgG seropositivity for that antigen. IgM antibody dilutions are shown as asterisks. The black horizontal dashed lines represent the upper and lower limits of the assay. **(B)** Measured IgG antibody dilutions and model predictions for the full population. Measured IgG antibody dilutions are shown as geometric mean titre (GMT) with 95% ranges. **(C)** Model predicted proportion of individuals testing seropositive over time.

Appendix Table 5: Parameter estimates for antibody kinetics model fitted to France data. Parameters of the antibody kinetics model are presented as posterior medians with 95% credible intervals. The model is fitted in a mixed-effects framework, so for every parameter we estimate the distribution within the entire population rather than a fixed value. We present the mean and standard deviation as summary statistics for the estimated distributions

description	parameter	prior	S^{tri}	RBD_{v1}	RBD_{v2}	S1	S2	NP_{v1}	NP_{v2}
mean of population-level distribution									
background IgG level	A_{bg}	0·001 (1·1x10^{-6}, 1·1)	2·8x10^{-5} (2·6x10^{-5}, 3·0x10^{-5})	2·1x10^{-5} (2·0x10^{-5}, 2·3x10^{-5})	4·2x10^{-5} (3·8x10^{-5}, 4·7x10^{-5})	5·2x10^{-5} (4·7x10^{-5}, 5·9x10^{-5})	4·5x10^{-5} (4·0x10^{-5}, 5·0x10^{-5})	5·4x10^{-5} (4·5x10^{-5}, 6·5x10^{-5})	8·1x10^{-5} (7·1x10^{-5}, 9·4x10^{-5})
ASC boost	β	0·01 (0·0001, 1·2)	0·00014 (9·7x10^{-5}, 0·0002)	3·9x10^{-5} (2·6x10^{-5}, 6·4x10^{-5})	8·2x10^{-5} (5·5x10^{-5}, 0·00012)	0·00020 (0·00012, 0·00032)	0·00017 (0·00012, 0·00026)	0·0040 (0·0013, 0·016)	0·0049 (0·0014, 0·023)
delay in generation of antibody response (days)	δ	5·4 (2·5, 15·1)	8·0 (6·0, 9·8)	9·8 (7·6, 11·5)	6·4 (4·1, 8·5)	9·3 (6·7, 11·3)	9·7 (7·7, 11·4)	8·4 (5·8, 10·6)	8·5 (5·6, 10·8)
half-life of memory cells (days)	d_b	2·1 (1·5, 4.0)	1·8 (1·4, 2·3)	2·2 (1·6, 4·0)	1·8 (1·4, 2·3)	2·0 (1·5, 2·8)	1·8 (1·5, 2·4)	2·4 (1·6, 15·5)	2·5 (1·6, 28·7)
half-life of short-lived ASCs (days)	d_s	3·2 (1·9, 9·2)	2·9 (2·1, 4·2)	3·0 (2·2, 4·3)	2·8 (2·1, 4·0)	3·0 (2·2, 4·3)	3·1 (2·2, 4·5)	3·0 (2·2, 4·2)	3·0 (2·2, 4·3)
half-life of long-lived ASCs (days)	d_l	400 (302, 567)	411 (229, 752)	416 (233, 745)	410 (228, 752)	413 (228, 746)	415 (233, 769)	416 (226, 775)	406 (223, 739)
half-life of IgG molecules (days)	d_a	21 (18·7, 24·1)	21·3 (18·9, 24·0)	21·3 (18·9, 23·9)	21·3 (18·8, 24·0)	21·4 (18·9, 24·1)	21·4 (19·0, 24·1)	21·2 (18·8, 23·9)	21·2 (18·8, 24·0)
proportion of short-lived ASCs	ρ	90% (65%, 95%)	72% (56%, 92%)	85% (0%, 94%)	75% (57%, 93%)	78% (56%, 93%)	70% (55%, 92%)	80% (57%, 94%)	81% (0%, 94%)
standard deviation of population-level distribution									
background IgG level	A_{bg}	0·0006 (6x10^{-7}, 0·8)	1·3x10^{-5} (1·1x10^{-5}, 1·6x10^{-5})	8·7x10^{-6} (7·5x10^{-6}, 1·0x10^{-5})	2·8x10^{-5} (2·3x10^{-5}, 3·5x10^{-5})	3·2x10^{-5} (2·6x10^{-5}, 4·1x10^{-5})	2·9x10^{-5} (2·3x10^{-5}, 3·7x10^{-5})	4·2x10^{-5} (3·0x10^{-5}, 5·9x10^{-5})	5·5x10^{-5} (4·2x10^{-5}, 7·4x10^{-5})
ASC boost	β	0·006 (5·4x10^{-5}, 0.9)	0·00014 (8·8x10^{-5}, 0.0002)	5·6x10^{-5} (2·8x10^{-5}, 0.00012)	8·5x10^{-5} (4·8x10^{-5}, 3·5x10^{-5})	0·00034 (0·00017, 0·00076)	0·0002 (0·0001, 0·0004)	0·04 (0·004, 0·53)	0·073 (0·0064, 1·13)
delay in generation of antibody response (days)	δ	3·5 (1·2, 34·6)	3·4 (2·5, 4·5)	3·3 (2·5, 4·3)	2·6 (1·6, 3·9)	3·2 (2·4, 4·2)	3·3 (2·6, 4·3)	2·7 (1·9, 3·8)	2·8 (1·9, 3·9)
half-life of memory cells (days)	d_b	1·1 (0·5, 7·2)	1·1 (0·6, 2·0)	2·2 (0·9, 9·2)	1·1 (0·7, 2·2)	1·5 (0·8, 4·0)	1·2 (0·7, 2·4)	2·9 (0·9, 162)	3·2 (0·9, 533)
half-life of short-lived ASCs (days)	d_s	2·3 (0·9, 29·2)	1·0 (0·6, 2·0)	1·1 (0·6, 2·2)	1·0 (0·6, 2·2)	1·1 (0·6, 2·1)	1·1 (0·6, 2·2)	1·1 (0·6, 2·1)	1·1 (0·6, 2·1)
half-life of long-lived ASCs (days)	d_l	109 (56, 349)	111 (46, 384)	113 (48, 371)	111 (47, 396)	112 (47, 377)	113 (46, 392)	113 (46, 394)	109 (45, 359)
half-life of IgG molecules (days)	d_a	3·2 (1·8, 8·6)	3·2 (1.8, 7)	3·2 (1·8, 7·2)	3·2 (1·8, 7·1)	3·2 (1.8, 7·1)	3·2 (1·8, 6·9)	3·2 (1·8, 7·0)	3·2 (1·8, 7·0)
proportion of short-lived ASCs	ρ	0·07 (0·02, 0·40)	0·30 (0·015, 0·45)	0·01 (0·0, 0·43)	0·28 (0·012, 0·45)	0·19 (0·0035, 0·44)	0·32 (0·02, 0·46)	0·19 (0·003, 0·45)	0·06 (0·0, 0·45)
observational variance									
log scale standard deviation for Luminex measurements	σ_{obs}	0·71 (0·18, 2·75)	0·47 (0·42, 0·53)	0·53 (0·47, 0·58)	0·59 (0·51, 0·66)	0·67 (0·60, 0·75)	0·67 (0·60, 0·75)	1·11 (0·98, 1·26)	0·91 (0·81, 1·02)

2.5. Model predicted diagnostic sensitivity over the first year of follow-up.

For each of the 215 individuals sampled, the mathematical model of antibody kinetics was used to simulate estimates of antibody levels during the first year of follow-up. To this data, we applied monoplex and multiplex classification algorithms to assess diagnostic sensitivity within the first year of follow up. These results are presented in Figure 3 (main manuscript) and in Appendix Table 6.

Appendix Table 6: Model predicted reductions in antibody levels and sensitivity at 6 and 12 months after symptom onset. All results are shown for a high specificity target (>99%). Estimates are presented as posterior median with 95% credible intervals.

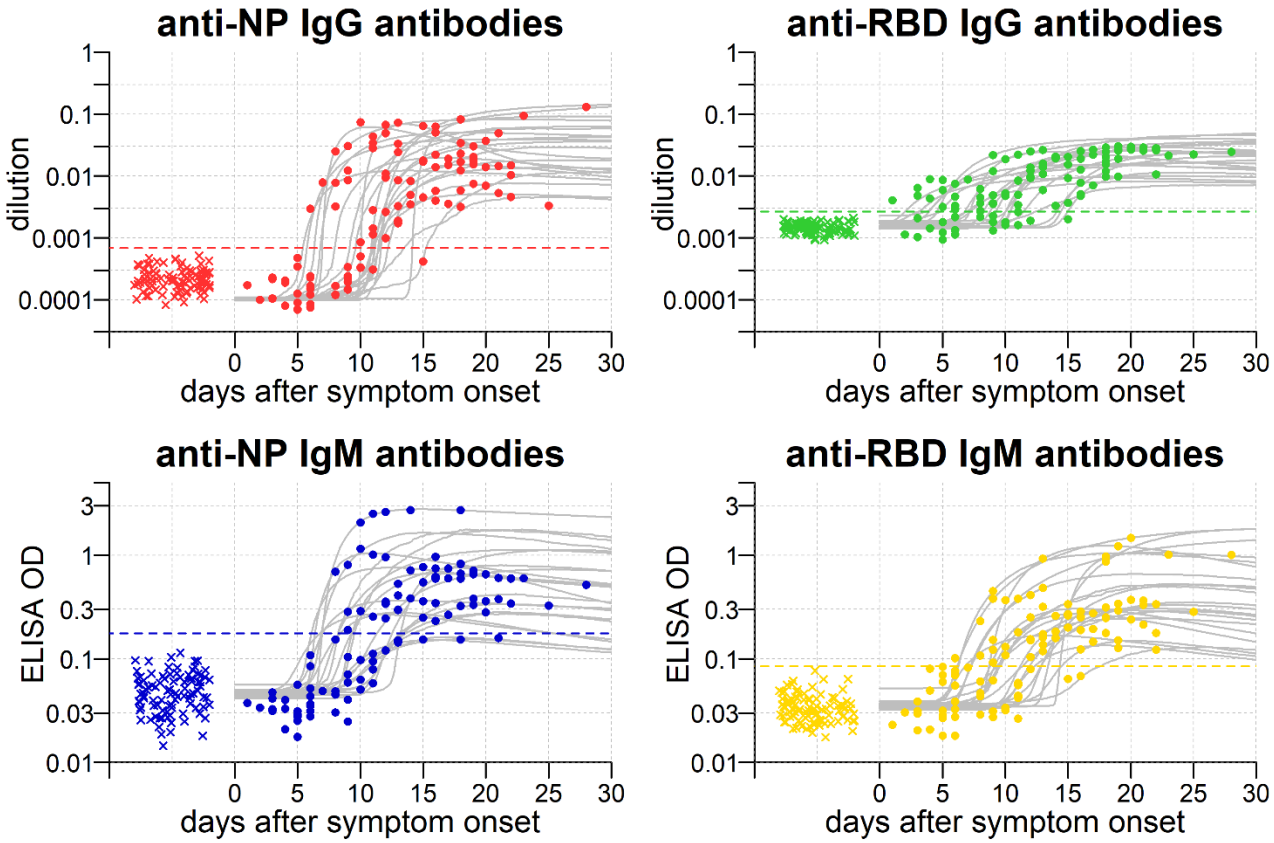
	model predicted reduction in antibody levels		**cutoff antibody dilution**	**model predicted sensitivity**	
	6 months	**12 months**		**6 months**	**12 months**
Single biomarkers					
S^{tri} IgG	23·3% (3·6%, 85·4%)	47·1% (17·5%, 90·3%)	0·000122	94·8% (71·4%, 99·2%)	88·7% (63·4%, 97·4%)
RBD_{v1} IgG	22·1% (1·9%, 94·8%)	44·3% (13·9%, 96·0%)	0·000024	89·7% (74·2%, 99·2%)	86·9% (70·1%, 98·7%)
RBD_{v2} IgG	23·2% (3·4%, 80·4%)	44·6% (16·1%, 85·8%)	0·000226	67·8% (46·6%, 80·7%)	56·7% (37·9%, 73·2%)
S1 IgG	20·5% (2·6%, 89·9%)	44·0% (14·6%, 92·4%)	0·000700	49·2% (31·7%, 67·6%)	41·2% (25·5%, 59·8%)
S2 IgG	18·3% (2·1%, 79·5%)	41·5% (13·5%, 86·5%)	0·000479	70·9% (51·0%, 87·7%)	61·6% (39·9%, 81·0%)
NP_{v1} IgG	27·4% (2·5%, 94·4%)	49·4% (11·5%, 96·4%)	0·000664	69·0% (53·6%, 78·4%)	63·9% (47·7%, 73·2%)
NP_{v2} IgG	25·5% (1·6%, 97·2%)	47·5% (7·8%, 98·4%)	0·000630	65·5% (49·0%, 77·8%)	57·7% (43·8%, 74·3%)
Multiplex combinations					
S^{tri} IgG + RBD_{v2} IgG	–	–	–	97·9% (89·5%, 100·0%)	95·4% (82·3%, 99·6%)
S^{tri} IgG + RBD_{v2} IgG + NP_{v1} IgG	–	–	–	98·4% (91·2%, 100·0%)	95·9% (83·9%, 99·6%)
S^{tri} IgG + RBD_{v2} IgG + NP_{v1} IgG + S2 IgG	–	–	–	98·9% (86·1%, 100·0%)	96·4% (80·9%, 100·0%)

3. Antibody kinetics: case study of Hong Kong data

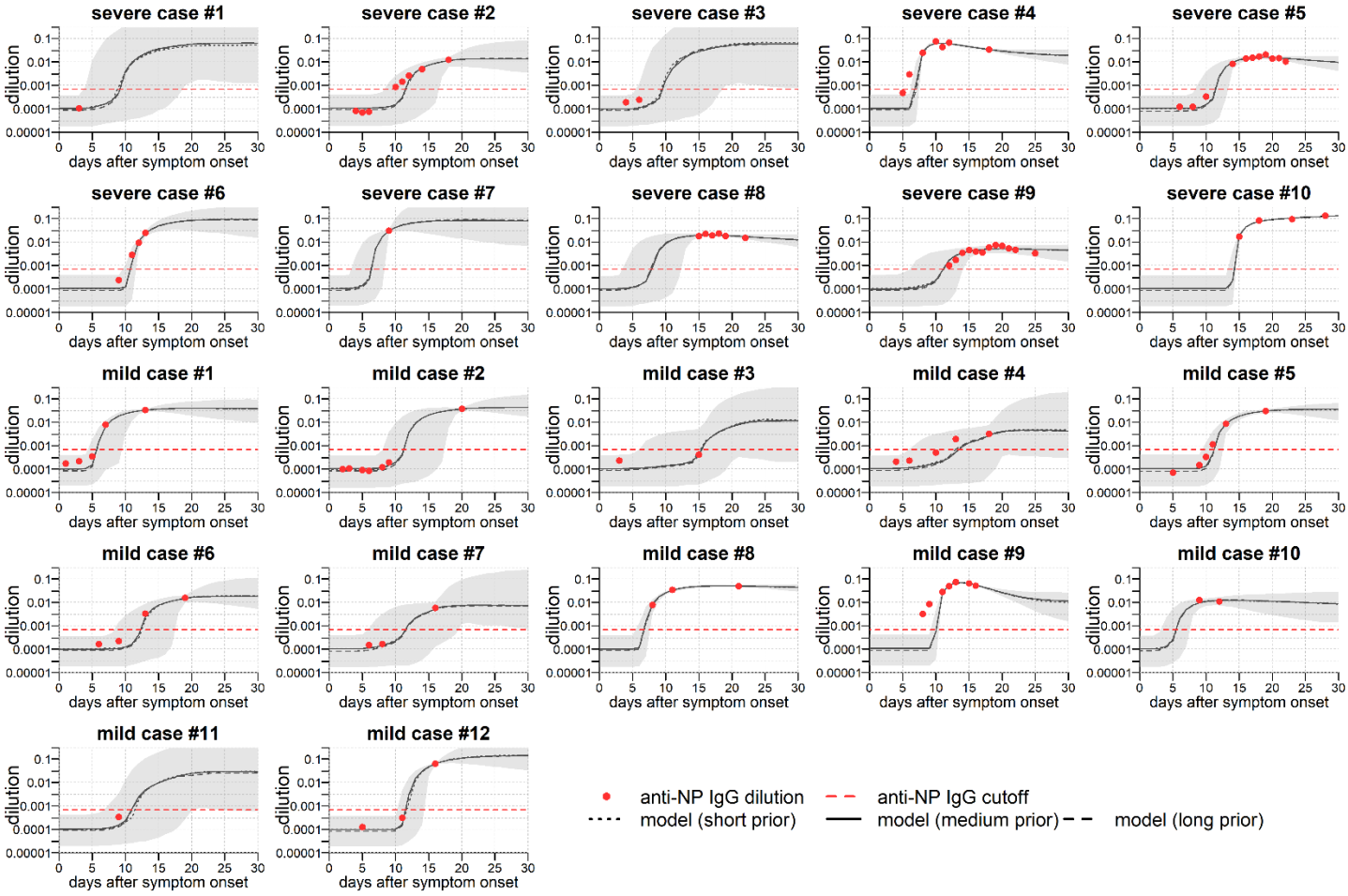
In order to obtain prior estimates of the parameters describing the early stage kinetics of the antibody response following SARS-CoV-2 infection, we performed a secondary analysis of data from patients admitted to Princess Margaret Hospital and Queen Mary Hospital in Hong Kong, following the primary analysis by To, Tsang *et al* [20]. 23 patients with RT-qPCR confirmed SARS-CoV-2 infection were followed longitudinally for up to four weeks after initial onset of symptoms. Ten patients had severe COVID-19, all of whom required oxygen supplementation, and 13 patients had mild disease.

The Hong Kong based team expressed and purified recombinant proteins for receptor-binding domain (RBD) and nucleoprotein (NP). Genes encoding the spike RBD (amino acid residues 306 to 543 of the spike protein) and full length NP of SARS-CoV-2 were codon-optimized, synthesized and cloned. IgG and IgM antibody responses were quantified via the optical density (OD) from an enzyme immunoassay (EIA). Serial dilutions from 1:100 to 1:16,000 of a positive control serum were assayed for IgG responses. This allowed conversion of IgG antibody responses measured by EIA OD to dilutions. To determine the sero-positivity cutoff, the mean value of 93 anonymous archived serum specimens from 2018 plus 3 standard deviations was used. The cutoff values were: anti-NP IgG = 0·523 OD; anti-RBD IgG = 0·108 OD; anti-NP IgM = 0·177 OD; and anti-RBD IgM = 0·085. After conversion of the EIA OD values to dilutions, the sero-positivity cutoffs for IgG antibody responses were anti-NP IgG = 0·00682; and anti-RBD IgG = 0·002665.

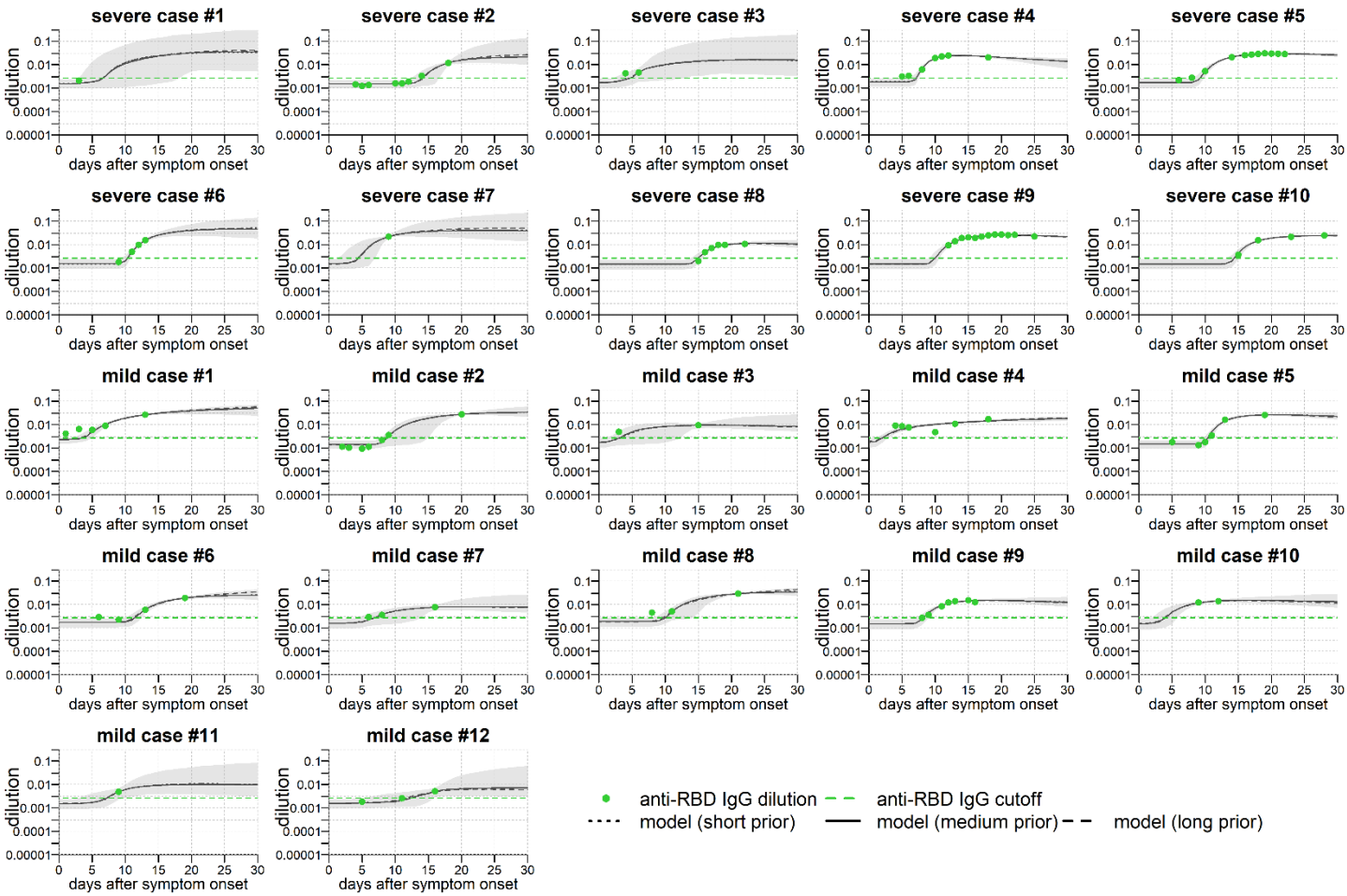
Estimated model parameters are presented in Appendix Table 7. Appendix Figure 6 provides an overview of the fitted antibody kinetics to all participants. Detailed individual-level fits to the data, with quantification of uncertainty are shown in Appendix Figures 6-9. Comparing the early kinetics of the IgG and IgM response, we estimate that the time to anti-NP IgG sero-conversion was 11·0 days (inter-quartile range (IQR): 8·1, 11·6), and the time to anti-NP IgM sero-conversion was 11·9 days (IQR: 8·4, 15·8). The time to anti-RBD IgG sero-conversion was 8·6 days (IQR: 5·3, 10·4), and the time to anti-NP IgM sero-conversion was 11·6 days (IQR: 9·2, 28·6). Although time to sero-conversion is dependent on the selection of sero-positivity cutoff, this suggests that IgM responses are not induced before IgG responses, and that both are generated at approximately the same time.



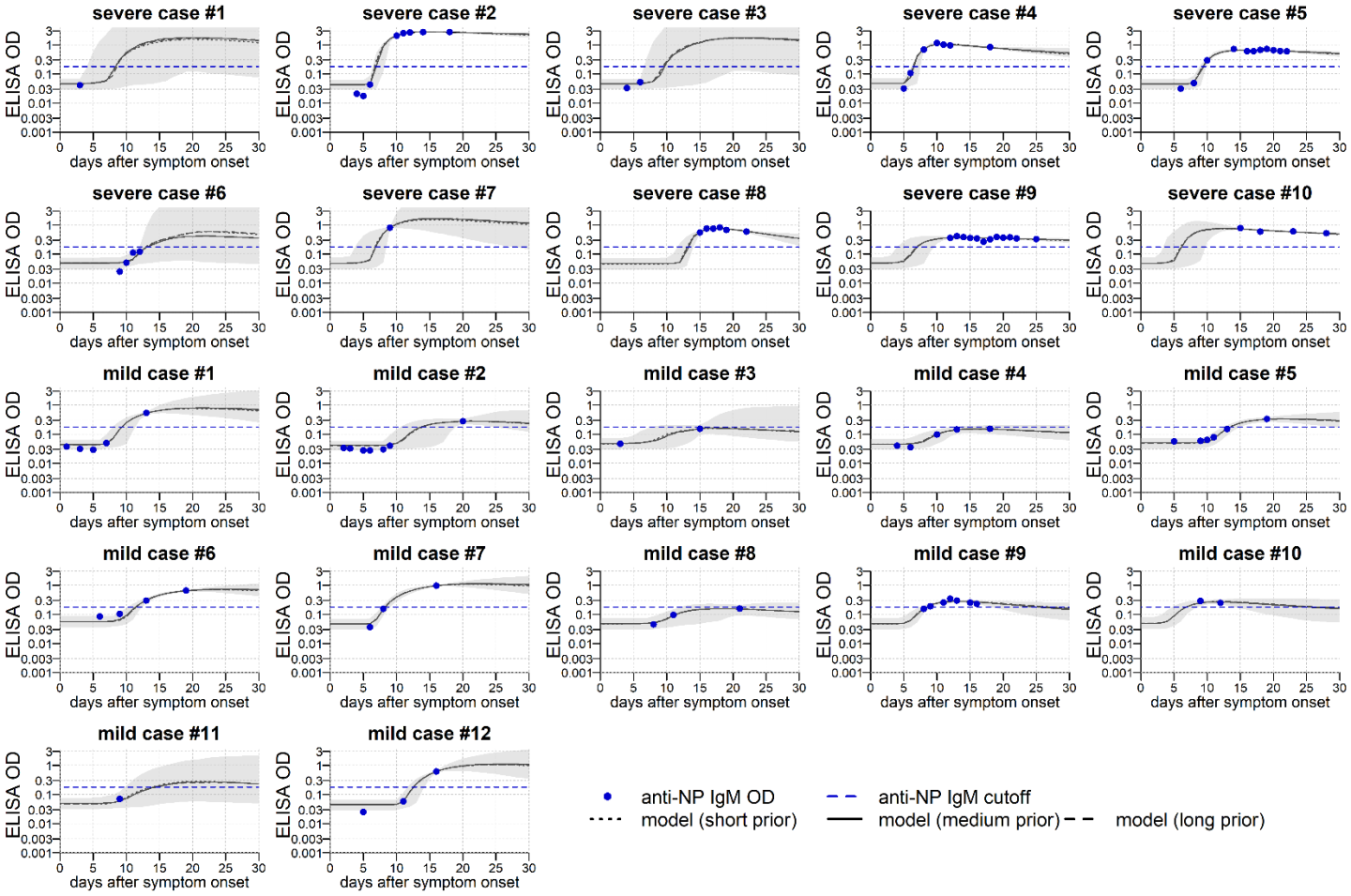
Appendix Figure 6: SARS-CoV-2 antibody kinetics in Hong Kong patients. Anti-nucleoprotein (NP) and anti-receptor-binding domain (RBD) antibody responses in 22 patients with PCR confirmed SARS-CoV-2 infection admitted to hospitals in Hong Kong. Measured antibody levels in patients are depicted as points. Measured antibody levels in negative controls are depicted as crosses. Grey lines show posterior median model prediction. The uncertainty of the model predictions is presented via 95% credible intervals in Appendix Figures 7-10. The horizontal dashed line represents the cutoff for sero-positivity.



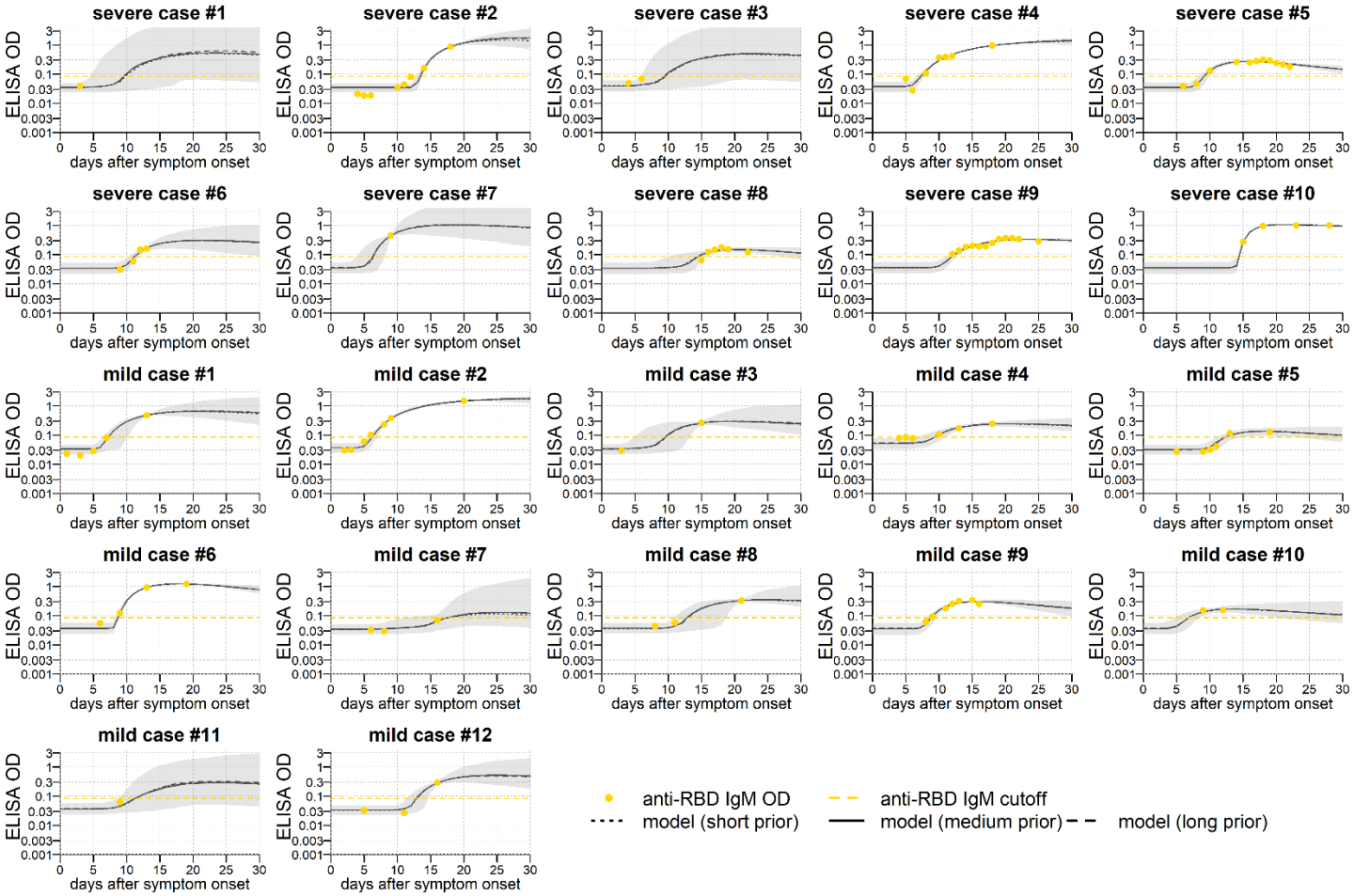
Appendix Figure 7: Model fit to short-term data on anti-NP IgG antibody responses. Measured antibody responses are shown as red points. Posterior median model predictions are shown as black lines, with 95% credible intervals in grey. The horizontal dashed line represents the cutoff for sero-positivity. Note that as there is no data on the long-term antibody response to SARS-CoV-2, three different sources of prior information were utilized. The half-life of the long-lived component of the antibody responses was assumed to be 200 days (short prior), 400 days (medium prior), or 800 days (long prior). Note that each of the three assumptions give near identical fits for the short-term kinetics displayed here.



Appendix Figure 8: Model fit to short-term data on anti-RBD IgG antibody responses. Measured antibody responses are shown as red points. Posterior median model predictions are shown as black lines, with 95% credible intervals in grey. The horizontal dashed line represents the cutoff for sero-positivity. Note that as there is no data on the long-term antibody response to SARS-CoV-2, three different sources of prior information were utilized. The half-life of the long-lived component of the antibody responses was assumed to be 200 days (short prior), 400 days (medium prior), or 800 days (long prior). Note that each of the three assumptions give near identical fits for the short-term kinetics displayed here.



Appendix Figure 9: Model fit to short-term data on anti-NP IgM antibody responses. Measured antibody responses are shown as red points. Posterior median model predictions are shown as black lines, with 95% credible intervals in grey. The horizontal dashed line represents the cutoff for sero-positivity. Note that as there is no data on the long-term antibody response to SARS-CoV-2, three different sources of prior information were utilized. The half-life of the long-lived component of the antibody responses was assumed to be 50 days (short prior), 100 days (medium prior), or 200 days (long prior). Note that each of the three assumptions give near identical fits for the short-term kinetics displayed here.



Appendix Figure 10: Model fit to short-term data on anti-RBD IgM antibody responses. Measured antibody responses are shown as red points. Posterior median model predictions are shown as black lines, with 95% credible intervals in grey. The horizontal dashed line represents the cutoff for sero-positivity. Note that as there is no data on the long-term antibody response to SARS-CoV-2, three different sources of prior information were utilized. The half-life of the long-lived component of the antibody responses was assumed to be 50 days (short prior), 100 days (medium prior), or 200 days (long prior). Note that each of the three assumptions give near identical fits for the short-term kinetics displayed here.

Appendix Table 7: Parameter estimates for antibody kinetics model fitted to Hong Kong data. Parameters of the antibody kinetics model are presented as posterior medians with 95% credible intervals. The model is fitted in a mixed-effects framework, so for every parameter we estimate the distribution within the entire population rather than a fixed value. We present the mean and standard deviation as summary statistics for the estimated distributions

description	parameter	prior	NP IgG	RBD IgG	NP IgM	RBD IgM
mean of population-level distribution						
background IgG level	A_{bg}	0·001 (1·1x10^{-6}, 1·1)	0·00011 (2·6x10^{-5}, 0·0003)	0·0015 (0·0013, 0·0017)	–	–
background IgM level	A_{bg}	0·03 (0·001, 1·0)	–	–	0·049 (0·043, 0·054)	0·036 (0·032, 0·04)
ASC boost in mild cases (IgG)	β_{mild}	0·01 (0·0001, 1·2)	0·014 (0·006, 0·051)	0·0028 (0·0015, 0·0053)	–	–
ASC boost in mild cases (IgM)	β_{mild}	0·11 (0·01, 1·2)	–	–	0·085 (0·048, 0·17)	0·08 (0·04, 0·16)
ASC boost in severe cases (IgG)	β_{sev}	0·01 (0·0001, 1·2)	0·028 (0·01, 0·207)	0·0056 (0·0034, 0·0099)	–	–
ASC boost in severe cases (IgM)	β_{sev}	0·11 (0·01, 1·2)	–	–	0·67 (0·29, 2·8)	0·14 (0·07, 0·46)
delay in generation of antibody response (days)	δ	5·4 (2·5, 15·1)	9·6 (7·7, 11·9)	7·8 (5·6, 11·7)	7·9 (6·4, 9·8)	8·7 (7·0, 10·7)
half-life of memory cells (days)	d_b	2·1 (1·5, 4·0)	2·0 (1·3, 7·8)	1·8 (1·3, 2·8)	2·0 (1·3, 5·7)	2·2 (1·5, 4·9)
half-life of short-lived ASCs (days)	d_s	3·2 (1·9, 9·2)	2·5 (1·8, 4·1)	2·4 (1·8, 3·7)	2·4 (1·7, 3·8)	2·8 (2·0, 4·7)
half-life of long-lived ASCs (days) (IgG)	d_l	400 (302, 567)	408 (227, 727)	417 (230, 771)	–	–
half-life of long-lived ASCs (days) (IgM)	d_l	100 (76, 142)	–	–	104 (68, 163)	103 (66, 167)
half-life of IgG molecules (days)	d_a	21 (18·7, 24·1)	43·5 (25·7, 243·6)	21·3 (18·4, 28·7)	–	–
half-life of IgM molecules (days)	d_a	10 (9.1, 11.5)	–	–	10·8 (9.3, 164.2)	10·2 (9.2, 13.2)
proportion of short-lived ASCs	ρ	90% (65%, 95%)	90% (79%, 94%)	80% (57%, 94%)	93% (65%, 97%)	89% (62%, 98%)
standard deviation of population-level distribution						
background IgG level	A_{bg}	0·0006 (6x10^{-7}, 0·8)	5·7x10^{-5} (1·0x10^{-5}, 0·00013)	0·0004 (0·0003, 0·0005)	–	–
background IgM level	A_{bg}	0·01 (0·0003, 0·5)	–	–	0·01 (0·007, 0·015)	0·008 (0·006, 0·011)
ASC boost in mild cases (IgG)	β_{mild}	0·006 (5·4x10^{-5}, 0·9)	0·020 (0·006, 0·23)	0·0017 (0·0007, 0·005)	–	–
ASC boost in mild cases (IgM)	β_{mild}	0·06 (0·004, 1·1)	–	–	0·045 (0·020, 0·17)	0·06 (0·03, 0·21)
ASC boost in severe cases (IgG)	β_{sev}	0·006 (5·4x10^{-5}, 0·9)	0·048 (0·01, 2·0)	0·0030 (0·0015, 0·008)	–	–
ASC boost in severe cases (IgM)	β_{sev}	0·06 (0·004, 1·1)	–	–	0·55 (0·19, 4·9)	0·17 (0·06, 1·4)
delay in generation of antibody response (days)	δ	3·5 (1·2, 34·6)	4·2 (2·8, 6·9)	6·5 (3·8, 17·5)	3·5 (2·3, 5·9)	3·8 (2·6, 6·4)
half-life of memory cells (days)	d_b	1·1 (0·5, 7·2)	1·8 (0·6, 35·3)	1·0 (0·5, 3·5)	1·8 (0·6, 18·5)	1·8 (0·7, 11·87)
half-life of short-lived ASCs (days)	d_s	2·3 (0·9, 29·2)	1·3 (0·6, 3·5)	1·2 (0·6, 2·8)	1·2 (0·6, 3·1)	1·6 (0·7, 4·6)
half-life of long-lived ASCs (days) (IgG)	d_l	109 (56, 349)	111 (47, 384)	114 (47, 404)	–	–

half-life of long-lived ASCs (days) (IgM)	d_l	22 (10, 69)	–	–	22 (11, 67)	23 (11, 73)
half-life of IgG molecules (days)	d_a	3·2 (1·8, 8·6)	84·0 (22, 2808)	5·4 (2·0, 27)	–	–
half-life of IgM molecules (days)	d_a	2·2 (1·2, 6·2)	–	–	4·4 (1·5, 2770)	2·6 (1·3, 12)
proportion of short-lived ASCs	ρ	0·07 (0·02, 0·40)	0·06 (0·02, 0·26)	0·25 (0·04, 0·45)	0·08 (0·02, 0·42)	0. ·18 (0·02, 0·44)
observational variance						
standard deviation for ELISA measurements (IgG)	σ_{obs}	0·004 (0·0002, 0·1)	0·0026 (0·0023, 0·0030)	0·0011 (0·0009, 0·0013)	–	–
standard deviation for ELISA measurements (IgM)	σ_{obs}	0·04 (0·002, 1)	–	–	0·031 (0·025, 0·037)	0·022 (0·019, 0·025)

4. References

1. Kirchdoerfer RN, Wang N, Pallesen J, Wrapp D, Turner HL, Cottrell CA, *et al*. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci Rep*. 2018; 8(1):15701
2. Longley RJ, França CT, White MT, Kumpitak C, Sa-Angchai P, Gruszczyk J, *et al.* Asymptomatic *Plasmodium vivax* infections induce robust IgG responses to multiple blood-stage proteins in a low-transmission region of western Thailand. *Malar J*. 2017; 16(1):178.
3. Fafi-Kremer S, Bruel T, Madec Y, Grant R, Tondeur L, Grzelak L, *et al*. Serologic responses to SARS-CoV-2 infection among hospital staff with mild disease in eastern France. *EBioMedicine*. 2020; 59: 102915
4. Grzelak L, Temmam S, Planchais C, Demeret C, Tondeur L, *et al*. A comparison of four serological assays for detecting anti–SARS-CoV-2 antibodies in human serum samples from different populations. *Sci Trans Med*. 2020; 12: eabc3103
5. Longley RJ, White MT, Takashima E, Brewster J, Morita M, Harbers M, *et al*. Development and validation of serological markers for detecting recent exposure to *Plasmodium vivax* infection. *Nature Med*. 2020; 26: 741-749
6. Wu LP, Wang NC, Chang YH, Tian XY, Na DY, Zhang LY, Zheng L, Lan T, Wang LF, Liang GD. Duration of antibody responses after severe acute respiratory syndrome. *Emerg Infect Dis.* 2007; 13(10): 1562-4.
7. Mo H, Zeng G, Ren X, Li H, Ke C, Tan Y, *et al.* Longitudinal profile of antibodies against SARS-coronavirus in SARS patients and their clinical significance. *Respirology*. 2006; 11(1):49-53.
8. Cao WC, Liu W, Zhang PH, Zhang F, Richardus JH. Disappearance of antibodies to SARS-associated coronavirus after recovery. *N Engl J Med*. 2007; 357(11):1162-3.
9. Liu W, Fontanet A, Zhang PH, Zhan L, Xin ZT, Baril L, *et al*. Two-year prospective study of the humoral immune response of patients with severe acute respiratory syndrome. *J Infect Dis*. 2006; 193:792–5
10. Tang F, Quan Y, Xin ZT, Wrammert J, Ma MJ, Lv H, *et al.* Lack of peripheral memory B cell responses in recovered patients with severe acute respiratory syndrome: a six-year follow-up study. *J Immunol*. 2011; 186 (12) 7264-7268
11. Callow KA, Parry HF, Sergeant M, Tyrrell DA. The time course of the immune response to experimental coronavirus infection of man. *Epidemiol Infect*. 1990; 105(2):435-46.
12. Choe PG, Perera RAPM, Park WB, Song KH, Bang JH, Kim ES, *et al*. MERS-CoV antibody responses 1 year after symptom onset, South Korea, 2015. *Emerg Infect Dis.* 2017; 23(7):1079-1084.
13. Mankarious S, Lee M, Fischer S, Pyun KH, Ochs HD, Oxelius VA, Wedgwood RJ. The half-lives of IgG subclasses and specific antibodies in patients with primary immunodeficiency who are receiving intravenously administered immunoglobulin. *J Lab Clin Med*. 1988; 112(5): 634-640
14. Janeway's Immunobiology, Ninth Edition. 2016. Garland Science: New York, New York. ISBN: (Paperback) 978-0815345053
15. White MT, Griffin JT, Akpogheneta O, *et al.* Dynamics of the antibody response to *Plasmodium falciparum* infection in African children. *J Infect Dis* 2014; 210: 1115–22.
16. Teunis PF, van Eijkeren JC, de Graaf WF, Marinović AB, Kretzschmar ME. Linking the seroresponse to infection to within-host heterogeneity in antibody production. *Epidemics* 2016; 16: 33–39.
17. Andraud M, Lejeune O, Musoro *et al*. Living on three time scales: the dynamics of plasma cell and antibody populations illustrated for hepatitis A virus. *PLoS Comput Biol* 2012; 8: e1002418.
18. White MT, Verity R, Griffin JT, *et al*. Immunogenicity of the RTS,S/AS01 malaria vaccine and implications for duration of vaccine efficacy: secondary analysis of data from a phase 3 randomised controlled trial. *Lancet Infect Dis* 2015; 15: 1450–58.
19. White MT, Bejon P, Olotu A, *et al.* A combined analysis of immunogenicity, antibody kinetics and vaccine efficacy from phase 2 trials of the RTS,S malaria vaccine. *BMC Med* 2014; 12: 117

20. To KK, Tsang OT, Leung WS, Tam AR, Wu TC, Lung DC, *et al*. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis*. 2020; S1473-3099(20)30196-1.